

FINDINGS

Studies of the implications of information technology for scholarly communication are diverse. The new technologies challenge the very definition of “publication” and raise new questions about the economic and legal underpinnings of the industry. The research is in its infancy and there is, as yet, little consensus except that the nature and role of scholarly journals are changing.

The literature is dominated by discussions of the relationship between the formal, peer-reviewed journal article and the larger hierarchy of scholarly and scientific communication modes and forms, and the extent to which the new information and communication technologies have altered and disrupted traditional roles and definitions. Some have seen the possibility for necessary reform of the publishing system, advancing with proposals that include eliminating or reducing the role of publishers; changing or eliminating peer review, which has historically been a function coordinated by the journal publishers; and changing how intellectual property rights are managed. Associated with this debate is the question of pricing electronic journals, which is part of a general discussion among economists about methods for pricing information goods and which has also become embedded in the concern among librarians over escalating prices for serials. Again, this represents an area of emerging research, and, as of this writing (January 2001), early experimental results are just beginning to be released. Models of pricing also require assumptions about how e-journals will be defined, used, and valued, which is an area in which results are still preliminary, diffused, and evolving as e-journals come to be more widely accepted.

This section briefly describes the scientific journal publication process; characterizes the literature; and summarizes the principal findings, methodological issues, and recommendations found in the literature as they relate to the five research questions on which our study focused, i.e.:

1. What issues arise from the literature?
2. How do information scientists measure “impact” or implications or effects?
3. Have changes in researchers’ behavior been discerned?
4. What are the implications for underserved populations in the United States or abroad?

5. Are information security (that is, how are systems and data protected from unauthorized use) and user privacy investigated?

The first question concerns issues that arise in the literature and parses into several topics, which are described in several subsections below: the changing nature of scientific communication and its implications for the role of scholarly journals, including the definition of an electronic journal; the character of the publication process; the interests and concerns of the various participants and stakeholders; and the attributes of the new e-journal article, or the new “artifact.” These imply changes in existing business models, intellectual property regimes, peer review, and archiving, all of which topics are discussed here.

The second and third questions concern evidence of changes in behavior and attitudes; they also parse into subtopics: How do researchers use e-journals in their own investigations?; this use is typically demonstrated through references in the published literature, i.e., citation analyses. How have behaviors and attitudes changed?; this is information typically captured through surveys and interviews. The literature relevant to both of these questions is discussed in the subsection titled “Measurements of Implications and Changes in Researcher Behavior.”

The fourth and fifth questions address the implications of electronic scientific journals for underserved populations and information security, including user privacy. The relevant literatures are discussed in the last two sections.

ISSUES IN THE LITERATURE: DEFINING ELECTRONIC SCIENTIFIC JOURNAL PUBLISHING

Studying electronic scientific journal publishing can be understood as the intersection of three important domains: (1) the sociology of science and scientific communication, (2) scholarly journal publishing (a formal process by which results are vetted and made available to the public), and (3) information technology. All three areas have rich intellectual traditions. As observed by

Alsop, Tompsett, and Wisdom (1997), “Within the scope of scholarly communities there are a number of stakeholders with different and conflicting interests...scholarly communication is so vast a topic that no single group can study it.” Griffith (1990), Pierce (1990), Lievrouw (1990), and Paisley (1990) offer excellent overviews of the development of the study of scientific communication, including discussions of the role of publication.¹²

Arms (2000, p. 276) describes an electronic journal as “an online publication that is organized like a traditional printed journal, either an online version of a printed journal or a journal that has only an online existence.” Kling and McKim (1999, p. 891; 2001, p. 8) have delineated distinctions among print journals, journals with print and electronic versions, and purely electronic journals. These last, in Kling and McKim’s typology, are accessed via electronic communication channels and include the concept of peer review; they are not to be confused or conflated with preprints, print journals delivered in electronic form, or nonpeer-reviewed journals.¹³

Scientific journal publishing may be differentiated from the broader field of scholarly journal publishing by domain, by the communities that form around the research, and by technical requirements. In print as well as digital, scientific journal publishing is rendered more expensive because of the use of equations, formulas, diagrams, figures, and images (Hitchcock, Carr, and Hall 1998b; King, McDonald, and Roderer 1980, p. 96; some of the technical issues are discussed in Entlich et al. 1997, pp. 108–15). The issue tends to be more acute in the physical sciences and mathematics.

Experiments in publishing scholarly journals in digital formats—distributed by CD-ROMS, for example—date to the late 1980s. Peek and Pomerantz (1998, pp. 330–37; see especially tables 1 and 2, pp. 332–33 and 334) provide a description of these early projects. Challenges at the time included converting the print archive of back issues and organizational structures based on the physical artifact. Factors pushing conversion included declining costs of computer-based storage, networking, increasingly ubiquitous personal computers, and good database

software (Arms 2000, p. 46). At roughly the same time that desktop computing and workstations became common among researchers, expansion of the World Wide Web after 1994 increased the momentum toward online distribution and storage, abetted by higher production costs that resulted in higher journal subscription prices (Varian 1997). Higher journal prices created an incentive to seek lower cost journals, presumably through electronic means.

The ability to search the full texts of articles, to search across multiple journals, and to browse online is consistently a desired feature of electronic journals. Eason et al. (1997), moreover, found differences between disciplines that sought electronic services, such as searching and browsing, and disciplines in which a new form of the artifact was appealing because the new artifact could support ancillary information. Since 1995, there has been intensified interest in various techniques for organizing, finding, storing, and displaying information in digital form generally grouped under the term “digital libraries.” This term can encompass a broad range of research activities, some of which are associated with traditional library institutions and others of which are more suited to the Web environment. The Digital Libraries Initiative jointly undertaken by the National Science Foundation, the Defense Advanced Research Projects Agency, and the National Aeronautics and Space Administration has been a focus for these activities, but by no means represents the totality of the research in this area. Indeed, various text and electronic publishing projects predate the initiative and have continued in parallel with it. Among the relevant technologies are those that support database construction, information retrieval, information searches, management of intellectual property rights, and systems for managing charges and payments.

ELECTRONIC JOURNAL PUBLISHING IN THE CONTEXT OF SCIENTIFIC COMMUNICATION

Numerous articles offer overviews that describe in varying levels of detail the structure of scientific journal publishing and the implications of the Internet/World Wide Web and other information technologies for scientific journal publishing (see sidebar). The studies describe the history of scholarly journals, the role of journals in scholarly communication, and the different cultures of scientific research.

¹²These essays predate the expansion of the Internet, although they all reflect the importance of databases of bibliographic information and full text articles.

¹³Such stringent distinctions are not universally made. Hahn (1998, p. 27) found, for example, in interviews with ecologists on their roles as authors that her respondents “referred to the electronic journal in somewhat monolithic terms” and that their views were relatively “uncomplicated.”

Background Resources

In addition to the context-setting materials discussed in this subsection, the literature also includes descriptions of early projects in creating finding aids such as Cambridge Scientific Abstracts (McGinty 1997), resources published by the Institute for Scientific Information (Kimberly 1995 and Trolley 1998), and pure online journals or electronic editions of print journals in:

- astronomy and astrophysics (Boyce 1997a, Van Steenberg 1994);
- chemistry (Bachrach, Burleigh, and Krassivine 1998; Entlich 1995; Entlich et al. 1995, 1997; Nitsche 1998);
- clinical trials (Hickey and Noreault 1992);
- computer science (Agosti, Crestani, and-Melucci 1998; Brueni et al. 1993; Denning and Rous 1997; Ensor and Wilson 1997; Moret 1997; Rous 1999);
- earth systems and sciences (Carr et al. 1997; Holoviyak and Seitter 1997; Ruzak et al. 2000);
- ecological science (Hahn 1998; Holling 2000);
- geology (Whalley et al. 1996);
- geophysics (Hedlund, Nechitailenko, and Sears 1998);
- library science (Widzinski 1997);
- mathematics (Birman 2000, Steinberger 1996, and Youngen 1996);
- physics (Bederson and Lustig 1997, Haynes 1999, Kelly 1997, and Singleton 1997); and
- radiology (Ackerman and Simonaitis 1997).

Many of these were projects undertaken jointly by Federal agencies, professional associations and societies, and university and commercial publishers. In addition, commercial publishers such as Academic Press (McKay 1996) and Elsevier Science (Hunter 1998a, 1998b, 1999a; Kluiters 1996; Tagler 1996a; Ter Meer and Zijlstra 1993) have described their activities in electronic journal publishing and in creating Internet-accessible collections of their publications. Although some of the technical issues raised in the early pieces no longer dominate the discussion (e.g., the treatment of graphics, CD-ROM versus online access), many of the concerns raised by these projects and many of the findings persist. Among these are attributes of good design, a concern for what users (both authors and readers) want, ownership of intellectual property rights (which affects the business model), speed and currency of publication, acceptance, and peer review.

Fjalbrant (1999); Oppenheim, Greenhalgh, and Rowland (2000); Schaffner (1994); and Tomney and Burton (1998) provide historical descriptions of the origins and evolution of scholarly journals. Piternick (1989) offers an interesting discussion of alternatives to the scientific journal prior to the expansion of the Web, noting deficiencies that include lag in publication; restrictions on length and illustrations; proliferation of journals; high costs to scientists and to scientific societies; and demands placed on volunteer editors, reviewers, and journal managers. Among the remedies proposed were radio/TV broadcast, distribution on tape and microfilm, delivery on demand, centralized depositories, archives for auxiliary material, and “preview” journals.

Lindholm-Romantschuk and Warner (1996) link monographs to the more general phenomenon of scholarly communication. In a citation analysis that covered philosophy, sociology, and economics, they identified persistent “core or elite groups” of individuals and publications in all aspects of the study. The two social sciences studied, sociology and economics, were characterized by greater reliance on the journal article than the monograph. This, the authors state, is roughly consistent with an earlier observation that the monographic literature tended to be more important in the humanities and social sciences than in the natural sciences where the journal article is preferred (pp. 394, 396).

One of the implications of the introduction of new communication technologies was thought to be a lessening of the elitist role played by traditional gatekeepers. Cohen (1996) studied this question, based on data collected in 1994 (i.e., before the full expansion of the Web), and found that faculty members who took advantage of new communication technologies published more frequently, had a higher incidence of prestige factors, and were more productive. He discerned no “democratizing effect” among the faculty. Within these conclusions, Cohen found that faculty used computer-mediated communication (CMC) to ask for assistance, exchange manuscripts, and interact with editors during the peer review process. Of concern were excessive e-mail, privacy, and intellectual property (Cohen 1996, p. 58).

Several studies (Alsop, Tompsett, and Wisdom 1997; Eason et al. 1997; Kling and McKim 2000; Till 2001; Walsh and Bayma 1996, 1997) examine the variation in communication within the sciences with direct implications for publication. Walsh and Bayma (1996, 1997)

explored the relationship between social context, specifically work organization, and the diffusion of computer network technology based on interviews in 1991–92 with 67 academic scientists in four disciplines (mathematics, physics, chemistry, and experimental biology), drawn at random from faculty lists for a private and a state university. They conclude that fields that consist of tightly coupled but geographically dispersed work groups (e.g., particle physics) adopt CMC systems more intensively than fields where work is performed within relatively autonomous groups (e.g., experimental biology). Other factors affecting adoption of the technology include institutional support (e.g., policies of the learned societies and publishers¹⁴), familiarity with and appropriateness of the technology (e.g., support for symbols), size and internationalization of the communities and work groups, and degree of commercial and market penetration. Thus, Walsh and Bayma note, chemists and experimental biologists (despite the existence of the databases of genetic sequence information) were reluctant to use informal methods of communication, including CMC modes.¹⁵ In the early 1990s, the dominant application was e-mail, and mathematicians, who were accustomed to working in isolation and communicating face-to-face at meetings, found informal CMC very useful as a means of reinforcing existing ties.

Both physics and mathematics had established preprints as a means of sharing information quickly, although formal publication was still required as a means of establishing the archival record and for promotion and tenure decisions (Walsh and Bayma 1997, p. 369). In contrast to the problems with journals, preprints (or technical reports) had long been a way for some communities of scientists to exchange papers after they were written but before they had been accepted for formal publication in scholarly, peer-reviewed journals. Preprint collections, particularly in physics but also in mathematics and computer science, were quickly migrated to the electronic environment (Canessa 1997; Goodman 2000; Kling and McKim 2000, pp. 2–3; and Walsh and Bayma

1997).¹⁶ Till (2001) reviews efforts in the 1960s to develop preprints in the biomedical sciences, which have not been early adopters of electronic preprint servers, and summarizes the characteristics of communities in which the electronic preprint technology was taken up quickly—most notably, the high energy physics community.

The two forms of communication—preprints and formal publication—were seen as compatible. Walsh and Bayma (1997, p. 378) quote an editor at the American Physical Society as saying that the preprint system, “works beautifully in a small community of people who work in harmony with each other and who know each other’s reputation.” The authors also note that the high acceptance rate (80 percent) reduced fears that the papers would not be “forthcoming journal articles” (p. 378). Youngen (1997) also comments that, although electronic preprints (e-prints) are becoming mainstream in many areas of physics and astronomy, the behavior of publishers has varied; some cooperate, while others refuse to accept manuscripts that previously appeared on the Internet/Web. Moreover, he continues, there are some areas where electronic preprints are unlikely to be accepted: where issues of patentability and proprietary information are involved (e.g., with regard to some engineering documents) and instances where widespread, informal distribution would be inappropriate “unless the item was being distributed for comment only.”

Odlyzko (1999) explicitly relies on the existence of the e-print archive model and carefully notes that his data are personal and observational and are most applicable to specialized journals with low circulation. Describing the physics e-print archive of the Los Alamos National Laboratory, he observes that the model “transforms the mode of operation of any community of scholars that embraces it” but then acknowledges that “it has been a struggle for pioneers in other areas to duplicate the process.” “Special cultural factors,” he continues, led to the quick adoption of the archive by the high energy physics community, but there remain many areas, especially in chemistry and medicine, where preprints are rare.¹⁷

¹⁴Note in this regard the decision of both the American Mathematical Society and the American Physical Society to adopt TeX, a markup language that handles equations, as a standard (Walsh and Bayma 1997).

¹⁵Recently, McMillan, Hamilton, and Deeds (2000) found that research-intensive firms that are open in supporting publication of their findings have higher R&D productivity than firms that have more restrictive policies.

¹⁶A variant on electronic preprint collections is Fast Track, the integration of two databases that covered patent applications in pharmaceuticals and agriculture. This resource for research chemists was seen as a vehicle for early dissemination of technical information (Gotkis 1992).

¹⁷Odlyzko has done careful work in developing cost estimates for publishers and libraries (see, e.g., Odlyzko 1997b).

Finally, Till (2001) compares adoption of preprints by the biomedical sciences and physics and identifies characteristics that account for the rapid adoption of electronic preprints among high energy physicists. These include the nature of the high energy physics community itself, which is well-defined and “voracious” in its reading; the high quality of papers in the physics preprint archive; the low level of concern in the field for patentability; a willingness to assign intellectual priority at “point of dissemination, rather than after peer review”; and—“the most crucial factor”—the extent to which “the author, reader, and reviewer communities have coincided,” contributing to a consensus about standards of quality for research (p. 11).

Two eLib (U.K.) projects considered the relationship between electronic journals and scholarly communication in the mid-1990s. Eason et al. (1997) examined 14 disciplines spanning chemistry and physics to history and law by combining surveys, interviews, and literature searches with two workshops. The researchers concluded that there was a wide variety of practice that did not conform neatly to distinctions between the natural and social (termed the “hard” and “soft” sciences in the study). Within a broad consensus that valued the access that electronic services (such as searching) offered, e-journals enjoyed differential acceptance based on the character of the research (predominantly solitary, e.g., mathematics, as opposed to team-based, e.g., chemistry); the importance of speed of publication; the prior role of the journal as the cumulative, authoritative record; the extent to which findings were concentrated in a group of journals or scattered among many journals; the ability of the electronic media to support other forms of data (particularly important to the chemists); the role of journals in teaching; and the relative views of authors and readers.

Alsop, Tompsett, and Wisdom (1997) reached similar findings in their examination of three eLib projects which spanned chemistry, law, and cultural studies. They emphasize the importance of existing communities and the extent to which these communities are information technology literate as preconditions to the adoption of electronic communication, including journals. They note that “lack of appropriate reward for electronic scholarly communication is a significant barrier to the development of electronic research communities.”

Kling and McKim (1999) lay out a series of distinctions among e-journals, which are accessed primarily in electronic form; hybrid p-e journals, which are primarily

available in print but also accessible online; and e-p journals, which are primarily distributed in electronic form but may have limited distribution in print. Among the sources of tension they identify are (1) efforts to reform scholarly publishing, (2) publishers’ policies with regard to prior publication and whether posting on the Internet represents prior publication, and (3) the extent to which posting an article to the Internet ensures the “appropriate audience” for it (p. 892). The authors then suggest that *effective* publication can be satisfied by three criteria: publicity, trustworthiness, and accessibility. They note that there are numerous ways that these criteria may be satisfied, that Web posting and effective scholarly communication are “loosely coupled” (p. 905), and that formal publication in journals—electronic otherwise—need not be affected by prior posting to the Web.

Smith (2000), describing initiatives at the American Physical Society, concludes that there is a continuum of communication forms of which the traditional, peer-reviewed scholarly journal remains an important feature along with private communication, technical reports, preprints, conferences, author websites, and review articles. Instead of consolidating all three critical functions of creation, communication, and criticism/confirmation in the scholarly journals, however, he suggests that, enabled by the Web, these journals “no longer form the primary communication medium.” Rather, they retain the important function of formal criticism and confirmation, which is critical in maintaining the integrity of the content of the preprint databases as well as contributing to personal research evaluations. Along the same lines, Heck (1996) calls attention to the need to recognize multiple kinds of electronic information forms, reflecting different communication functions. Sandewall (1997) makes a similar case, differentiating in the electronic environment between “first publication” archives and archives of articles that have been subject to peer review.

Kling and McKim (2000) address differences among fields and electronic communication within high energy physics, molecular biology, computer science, artificial intelligence, astrophysics, and information systems. They argue that a “highly configurable technology such as the World Wide Web can be adopted and used by different fields in dramatically different ways” (p. 3). Among the factors that contribute to differences among the fields are policies of the professional societies toward posting to the Web; notions of trustworthiness (e.g., particle physicists, in whose field there is a tradition of preprints, are more willing to use working papers than are molecular

biologists or sociologists, while high energy physicists and computer scientists are more comfortable with circulating working papers than are chemists, molecular biologists, and psychologists who prefer to read “peer reviewed articles when the topic is outside the envelope of their current expertise” [p.4]; research costs; visibility of ongoing work; industrial integration, particularly where there is ready income from patents and trade secrets; and concentration in communication channels (e.g., astrophysics relies on 3 to 4 journals while neurology relies on well over 100 [p. 5]). Thus, the authors speculate that although “some conventions of scientific communication” are likely to change in the next few decades, the “fields will differ in the ways that shape e-media because trust issues work out differently with respect to characteristics” such as visibility of projects to others in the field and industrial integration (p. 6).

In a third study (Kling, Fortuna et al. 2000), Kling and his colleagues look at the proposal to establish a centralized electronic archive of all biomedical research articles. The original proposal called for two sections: a preprint server to which authors could submit their own biomedical research papers that had not yet gone through traditional publication channels and a peer-reviewed published-document server. The proposal suggested that this new biomedical research archive, E-Biomed, would be “free access,” with no fees or credentials required to access the archive. The resulting service, PubMed Central, is quite different. Scientific societies and commercial publishers maintain central roles in content control and dissemination, determining what is put up on the archive and when it is posted.¹⁸

The authors examine the reasons for the transformation, including analysis of a listserv established to enable members of the communities affected by the proposal to voice their opinions. They conclude that the professional societies and commercial publishers were able to rebalance the service away from authors/readers for several reasons, including their ability to align the interests of their memberships with their positions (as opposed to seemingly lone voices on listservs); their access to the director of the National Institutes of Health (NIH), who led the E-Biomed effort; their access to the press and to

congressional funding sources; and their ability to control access to the published article, which was essential to the ultimate utility of the service. Comparisons with other listservs suggest that the interests and views of authors and readers change when the discussion moves from the abstractions of electronic publishing to the immediacy of a given proposal. Indeed, participation in the E-Biomed listserv was surprisingly low; only 224 people participated, 125 of whom were supportive of the original proposal. The scientific societies, however, represented thousands of members, and had differential direct access with the director of NIH. Finally, comparison with Los Alamos preprint server suggests to the investigators that differences lie in the networks of social ties between the members of the disciplines, the Federal Government, and the general public. These ties include, but are not limited to, the existence of a tradition of preprints, relationships with scholarly societies, and political and financial relationships with public funding agencies.

EXTENT OF ELECTRONIC JOURNAL PUBLISHING

Electronic journal publishing is growing rapidly. According to Rowland (1999, p. 209), the first newly founded, free electronic journals were established in about 1990.¹⁹ Launch of the World Wide Web led to the rapid proliferation of these journals from about 1992–93 onwards. After 1995, established commercial and non-profit publishers began to make journals available in parallel print and electronic forms. Hitchcock, Carr, and Hall (1998b) and Peek, Pomerantz, and Paling (1998) roughly concur in this assessment of the timing, although the latter researchers narrow the entry of the major academic publishing houses to 1997.²⁰ Luther (1997) also focuses on 1997, offering a description of the e-journal services provided by Academic Press, Blackwell, EBSCO

¹⁹Note, however, that electronic access to full-text articles did not originate with the establishment of electronic journals. Schauder (1994, p. 77) documents access to electronic texts through abstracting and indexing services or database hosts beginning in the 1970s with the establishment of LEXIS (for case law) and NEXIS (for business and news). In the 1980s, Elsevier and the American Chemical Society made material available through BRS, DIALOG, and other dial-up services. By 1992, Schauder reports, the classification “full text database” was recognized in the *Cuadra/Gale Directory of Online Databases*.

²⁰Scholarly publishing anticipated the move to online. Major commercial publishers Time-Warner, Simon & Schuster (a unit of Viacom), and Random House (a unit of Bertelsmann) announced major online/digital initiatives in the spring of 2000; see Caravajal (2000). The early adoption of the technology among STM journals is discussed by Arnold (1999).

¹⁸Judith Turner discusses the development of PubMed Central in *iMP magazine* in November 1999. *iMP* was a nonjuried online magazine published by SAIC and edited by the principal author of this report while she was at SAIC; the article is not discussed, given potential conflicts of interest, but it is herein reported for use by future researchers.

Information Services, Elsevier, HighWire Press, Information Quest, Institute of Physics, Johns Hopkins University Press, OCLC, Ovid, Springer, and SWETS.²¹

Because the numbers of e-journals are growing rapidly and because there is disagreement over what constitutes an e-journal, precise numbers are relatively hard to come by. The Association of Research Libraries (ARL) publishes the *ARL Directory of Electronic Journals, Newsletters and Academic Discussion Lists*. According to the most recent (1997) edition, there were 27 e-journals/zines (i.e., both peer and nonpeer-reviewed) in 1991; this number increased to 2,459 in 1997. Growth in peer-reviewed e-journals only was from 7 in 1991 to 1,049 in 1997.

Journals in science, technology, and medicine (STM) have tended to represent the leading edge in the adoption of electronic journals as a means of formal communication. Of the electronic journals identified by the *ARL Directory*, 34 percent are in the life sciences, physical sciences, and technology; an additional 28 percent are in the social sciences. When the definition is limited to “pure” electronic journals—that is, journals that only exist in electronic form—Kling and McKim (2001, p. 8) observe that “there are remarkably few pure electronic journals in the sciences.”

Tenopir and King (2000, p. 344) accept ARL’s data but note that by 1999, “an estimated 4,000 journal titles were available electronically” either directly from the publishers or through third-party aggregators, such as abstracting and indexing services (p. 344). In this regard, Wicks (1998, p. 147) reports that the Chemical Abstracts Service (CAS) covered 28 electronic only journals as of April 1998, the Institute of Scientific Information covered 17 hybrid and “pure” electronic journals, the National Library of Medicine covered 5 pure e-journals, BIOSIS covered 9 (as of May 1998), and PsychINFO projected possible coverage of 24 e-journal titles as of 1998 or 1999.

Hitchcock, Carr, and Hall (1997) review several sources of information and conclude that there were 25 peer-reviewed e-journals in science in 1994, 115 in STM in 1995, and about 1,300 peer-reviewed e-journals based on publishers in the United Kingdom in 1997 with a projected circulation of 3,200 subscriptions in the United

Kingdom in 1998–99 based on then-current growth projections. Based on 1990–95 data, Hitchcock, Carr, and Hall (1998b) note that STM electronic journals come from three main sources: commercial publishers, learned societies, and research institutions including universities. Resh (1998), citing data from Abate 1997, provides a slightly different breakdown: According to data compiled in the mid-1990s, online publishers of scientific information include commercial publishers (40 percent of total); nonprofit scientific societies (25 percent); and university presses, government, and private research institutions (35 percent).

THE E-JOURNAL PUBLICATION MODEL

King, McDonald, and Roderer (1980, p. 17) describe the conventional scholarly journal publication model, noting that “there are several generic functions that must be performed in a journal system: that there are four principal participants (scientists as authors and readers; publishers; libraries; and secondary organizations), and that the flow of information can be aggregated into a reasonable number of paths...” The researchers cite these functions as follows: (1) generation of knowledge, or research; (2) writing, editing, and reviewing, or composition; (3) recordation, i.e., recordation of a manuscript as a formal communication; (4) reproduction; (5) distribution; (6) acquisition and storage; (7) control, i.e., provision of access through indexing etc.; (8) identification and location by others; (9) physical access; and (10) assimilation by other users in new research (pp. 14–15). King retains this model in a subsequent study (Tenopir and King 2000, p. 91).

Twenty years ago, King, McDonald, and Roderer (1980) argued that authors had little impact on the scientific community by means of formal communication until the work was published. One of the implications of the work in scientific communication that has been previously described (e.g., Cohen 1996) is that e-mail and other forms of computer-mediated communication amplified the influence and scope of informal communication mechanisms. Within the framework of electronic journals, much of the discussion among scholars, which will be described in the next paragraphs, can be understood as confusion over and realignment of functions that have existed in print, as described by Tenopir and King (2000), exacerbated by rising costs for both publisher and libraries and allegations of excessively high prices and monopolistic behavior (Ekman 1996 and Oppenheim, Greenhalgh, and Rowland 2000, pp. 362–33; on functions, see also Sloan and Okerson 1994 and Watkinson 1998, p. 26).

²¹Walter (1998) speculates about the significance of brand identification among commercial publishers vying for users’ attention on the Internet. The timing of this piece is consistent with the apparent importance of the year 1997.

Rowland (1999) describes two basic views of the future of scholarly publishing. One, held by a group comprising commercial publishers as well as many learned society publishers, expects current commercial conditions based on the existing print model to survive. A second group argues that the new communications technologies enable “academics to regain control over their own communication system from the commercial publishers” (p. 210). Rowland notes that this view is put forth by several influential individuals and finds support among librarians as well. Kling and McKim (1999, p. 892) have characterized this latter perception as the focus of a “small e-publishing professional reform movement” with a core group of highly articulate enthusiasts who raise important issues but bias the discussion by claiming that “a single model for electronic scholarly publishing is appropriate for all scholarly communities” (p. 893).

The themes raised by Rowland, Kling, and McKim are echoed throughout the literature. Descriptions of commercial projects, consortial efforts, and proposals outlining various publishing models and strategies are described in Berry (2000); Birman (2000); Creth (1997); Luther (1997); Peek, Pomerantz, and Paling (1998); and Rowland (1999). There are also proposals that consider various noncommercial alternatives involving free e-journals, pre-print archives in lieu of and in parallel with juried publication, self-publication through individual websites, and publication by universities and libraries (see, e.g., Okerson 1996). Self-organizing approaches have been discussed by, among others, Harnad (1996); Odlyzko (1996a); Smith (1999); Stodolsky (1995); and Varian (1998). As Smith (1999) and Varian (1998) point out, in addition to mechanisms for quality control, there is a need for stable archiving of the record. Finally, Hitchcock, Carr, and Hall (1998a) offer four specific steps toward “optimum e-journals”: nonexclusive papers, archives and gateway services, open systems, and links.

Of all of these proposals, Varian’s (1998) is probably the most carefully developed, taking into account potential cost savings from reengineering the process as well as using threaded discussions²² as a means of reader-based evaluation. In addition, he considers the savings in terms of library shelf space, monitoring for use of journal

acquisition as well as promotion and tenure review, search, and access to related and supporting material. He also considers the implications of the “network externality effect” (when the value of a good—in this case, a journal—depends on how many other people use it).

ISSUES IN THE LITERATURE: STAKEHOLDER INTERESTS AND CONCERNS

The previous section considered the structure of the publication process and how it is changing; this section considers views of scientists, authors, publishers, and librarians. Scientists’ views of the functions of scholarly journals vary depending on whether they are authors, readers, reviewers, editors, or members of editorial and library boards—any of which roles they can concurrently or intermittently perform during their professional lives (Hunter 1999a, 1999b; Oppenheim, Greenhalgh, and Rowland 2000, p. 369; Watkinson 1998, p. 26). The following subsections examine the literature on scientists as authors and readers as well as presenting the interests and concerns depicted in the literature for two other major stakeholders in the scholarly journal arena: publishers and librarians. Additionally, information is presented about tensions among some of these stakeholders (see sidebar).

AUTHORS

Scholarly authors appear to be motivated by the desire to communicate with their peers; they also view publication as important to promotion and tenure. Finally, studies suggest that scholarly authors are generally satisfied with peer review although they can also be frustrated by the delays peer review can occasion. The Association of Learned and Professional Society Publishers supported a survey of 11,500 authors in the arts and sciences who had contributed to journals published by members of the association (Swan 1999). Based on 2,500 responses, “communication with peers” was cited as authors’ principal motivation for publishing, followed by “prestige” and “funding for future research.” Authors in the sciences tended to place greater importance on future funding than did authors in the arts. Authors selected the journals in which to publish based on their reputation, “impact factor,” and reach and coverage via abstracting and indexing services. Science authors were more concerned with the existence of electronic versions

²²Threaded discussion forums allow users “to share ideas using the Web’s hypertext capabilities. Discussion forums link messages by subject. Thus, all messages on one topic are grouped together, allowing users to follow connected *threads* of thought.” (Definition taken from the University of North Carolina–Greensboro, Division of Continual Learning, website, <http://www.uncc.edu/cex/common/discuss.htm>.)

Libraries Versus Publishers

As Rowland's (1999) discussion of electronic scholarly communication implies, there is a tension between publishers and libraries as well as between publishers and users; this is explored more fully later in this section in the material addressing pricing and business models, since the issue comes into sharp relief in the context of the pricing of serials. Historically, as King, McDonald, and Roderer (1980) have shown, both publishers and libraries have had roles as aggregators and intermediaries, and the literature from both communities has variously discussed threats and challenges posed by the new electronic communication technologies. However, it is argued that the new communications environment, in particular the Web or the combination of the Web and electronic preprint servers, obviates the need for publishers or greatly reduces their role, particularly if they are no longer required to coordinate peer review (see, for example, Harnad 1996; Hayes 1996; Odlyzko 1996a; Oppenheim, Greenhalgh, and Rowland 2000, p. 366; Smith 1999). Other observers argue for a greater role for society and university rather than commercial publishers (Creth 1997, Johnson 2000, Pikowsky 1997). This may be more of a concern for authors than for readers. In their interviews with members of departments of mathematics and science in the United Kingdom, Pedersen and Stockdale (1999, p. 49) found that readers, for the most part, were not concerned with who published the journals they used, "although a small percentage preferred professional society journals." More important for both readers and authors was the "quality and impact rating of the journals." The investigators did not find that users linked any type of publisher with quality.

One early project that looked directly at issues related to libraries and publishers was The University Licensing Program (TULIP), which started in early 1991 and concluded at the end of 1995 (Borghuis et al. 1996, Lynch 1995, Mostert 1995). The project involved Elsevier and nine university participants in the United States. The technology was overtaken by the widespread deployment of the Web, but many of the management and operational experiences remain technology independent. The project's user studies concluded that users of the system sought convenience, searchability, hyperlinks, sufficient journal coverage and timeliness, and speed—both in downloading and printing. Participants did not consistently conclude that the availability of material in electronic form necessarily eliminated either libraries or publishers but did report that cost was a factor—despite an initial belief that the digital technologies would reduce costs.

of journals, and chemists, in particular, were concerned over publication speed. Science writers were also more concerned over reproduction quality. Interestingly, authors in the social sciences and arts were more concerned over retaining copyright than were science writers, who, for their part, were more concerned that "someone else will publish very similar work first."

About two-thirds of the respondents were "generally satisfied" with peer review, but more than half still considered it "an obstacle in achieving their objectives when publishing work," citing delays occasioned by reviews, and superficial and unnecessarily hostile reviews.²³

Swan's findings are more or less consistent with those reported by Schauder (1994) and Tenopir and King (2000), who also find that communication with peers, contributing to knowledge, and career advancement are powerful motivators, based on research that spans the 1970s to early 1990s. Tenopir and King found concern over publication lag but concluded that electronic transmission of manuscripts was unlikely to reduce the delay since "delays occur while the manuscript awaits actions by editors and authors" (p. 156). In 1997–98 interviews with ecologists about their roles as authors and editors, Hahn (1998, p. 28) found some concern over the quality and stability of electronic journals, in which ease of publication was associated with the ease of publishing "junk" and, more generally, "perceived threats to the peer review process."

Studies have begun to probe how writing itself might change in the context of the electronic journal. Kneece (1996) looked at writing for electronic publishing and concluded that digital documents possess two properties: facilitating the process of "automated retrieval" and helping readers to decide quickly whether it would be worth "reading online, downloading, or printing" (p. 198). She assumed a fairly limited interface that would permit readers to see only a small part of the document, arguing that writers could assist readers by providing "visible structure" at the top of the document: a table of contents, summary, or list of bullets. Sonkkila (1998) argues that changes in the publication process as a result of the efficiencies and capabilities of the new information

²³In this regard, however, Watkinson (1998, p. 26) found that "there is absolutely no evidence that scientists want to give up their structured papers or the prerequisite of peer review" as they embrace electronic journals.

technologies to serve the needs of different research communities may place a greater burden on authors, who might be expected to undertake more complex mark-up than under the print regime. Bern and Meinel (1999) describe efforts at the department for Theoretical Computer Science at the University of Trier, Germany, to develop Internet-based methods and tools that enable scientist-authors to collaborate and publish their research.

Finally, Rowland (1999) suggests that a key issue for authors is intellectual property: Who should own the underlying copyright? It is unclear if scholarly authors uniformly resent having to convey copyright to the publisher. Schauder (1994, p. 92) reports from his 1992 survey of 743 senior academics that, when asked “whether they considered it fair that they had to cede copyright, the most common response (122 of 289) was ‘don’t know.’” (Schauder’s survey is discussed in more detail later in this report, as is the topic of intellectual property.) Associated with copyright is the question of whether the conventional publisher, profit or nonprofit, should obtain exclusive rights of distribution. Bachrach, Burleigh, and Krassivine (1998) make the case for the authors of government-supported research to “distribute these works as they see fit, via journals, electronic postings, and other new modes that may appear.”

READERS

One consistent finding across the studies considered in this review is the importance of information technology in making use of the research literature more convenient. In general, studies of what readers (or users) want in electronic journals (or systems) are numerous and are difficult to compare, since the questions, concerns, and variables employed differ from study to study. This literature review distinguishes between what readers want and attributes of the new electronic artifact, since the former reflects *behavior* while the latter reflects *properties* of a system or artifact. Properties of the artifact are discussed later in this section.

Tenopir and King (2000) report the results of studies conducted between 1984 and 1998 on how scientists in corporations, government agencies, and national laboratories spend their time. Most of their time (78 percent) was spent conducting research, but more than half of their time was spent in communications-related activities, especially reading. University scientists tend to read more than other scientists, but most readings of scientific scholarly articles are by nonacademic scientists “simply because nonacademic scientists greatly outnumber

academic scientists” (p. 132).²⁴ Tenopir and King provide detailed discussion of patterns of readership among academic and nonacademic scientists (see chapters 6, 7, and 8). They conclude by noting that the interests of readers and authors, while frequently closely aligned, are not identical, as is sometimes asserted. Björk and Turk (2000) report related findings from a survey done in 1998 by the International Council for Scientific and Technical Information. This study showed that 61 percent of the respondents felt that “electronic journals or trade magazines are easy to use or user friendly, but that only 14 percent of respondents publish in such journals.”

Oakeshott (1985) describes the BLEND (Birmingham and Loughborough Electronic Networking Development) project, an experiment in electronic communication funded by the British Library between 1980 and 1984. Technologically, the system required dedicated terminals and telephone lines; limited access to the system and the system’s relatively limited capabilities proved a deterrent to use. However, the experiment did yield interesting results on patterns in reading and information use. Reading habits varied from those who read articles completely before determining their relevance to those who browsed and scanned.²⁵ Many read outside conventional office hours, either at home or while traveling. Informal search methods (by librarians’ standards) were popular, and the readers preferred release of new issues on “set dates,” because it reduced the costs of multiple log-ins (p. 32). Additionally, users seemed to like the ability to print out electronic materials, “reassured to have a copy to hand should it ever be required.” Oakeshott (p. 35) concluded with a call for further research into “how people read, search, and write, as well as into the creation of text which can cope with the different types of use and user for input and output—both reading and writing.”

The importance of browsing and searching as a means of finding information and hence in contributing to the importance of online journals has been supported in subsequent studies (see Baldwin 1999, Entlich et al. 1996, Kirstein and Montasser-Kohsari 1996, Pedersen and Stockdale 1999, Stewart 1996, Tenopir and King 1998). Many of the technical limitations apparent in the BLEND project no longer exist, at least for many scientists in the United States. Moreover, since BLEND, user

²⁴The researchers note, however, that academic scientists appear to “write significantly more for external consumption than do other scientists” (Tenopir and King 2000, p. 131).

²⁵In this regard, Oakeshott notes that browsing could “take as many forms as the number of people involved.”

sophistication has advanced, and more information has become available electronically. There is some evidence that readers' "wish lists" have similarly expanded. Thus, Baldwin, reporting on the eLib SuperJournal project (see sidebar later in this section), found that "users want quick and easy access to a wide range of quality journals that are up to date, and guaranteed access to the journals once they have found them" (1999, p. 214). She found that key benefits for readers were convenience and time saving. Social scientists surveyed showed more interest in the range of journals and in backfiles; natural scientists showed greater interest in timeliness.

The second feature of the BLEND study to note is readers' desire to print out materials. The Chemistry Online Retrieval Experiment (CORE) was a five-year R&D project undertaken by the American Chemical Society, its Chemical Abstracts Service, Bellcore (now Telcordia, a subsidiary of SAIC), OCLC, and Cornell University (Entlich et al. 1996). In addition to the importance of searching and convenience, the CORE team found that one of the values of an electronic collection was that users tended to find relevant material in unanticipated sources (p. 111). Although readers consistently preferred to discover information via online sources, they printed them out in order to read in-depth (p. 110). Stewart (1996, p. 341) concurs with this finding, reporting that the single most important feature for users was the ability to print. Jasperse and Hawcroft (1992), Kirstein and Motasser-Kohsari (1996, p. 94) and Schauder (1994, p. 91) reported similar results.

Nuclear Technology Publishing (NTP) conducted a survey of 20,000 contacts in its database as part of an effort to establish online requirements for its readers. The survey was distributed as part of a mailing as well as via the publisher's website; in all, 500 individuals responded for a response rate of approximately 2.5 percent, which was considered sufficient by Goldfinch (2000, p. 243) to support analysis. Of those who responded via the print survey, 57.7 percent accessed information via the website, compared to 79.2 percent of those who responded via the Web-based survey. Slightly more than half of all respondents felt that it was very important to have current year abstracts accessible online. Less interest was expressed in having back issues accessible via the Web; CD-ROM was considered a viable distribution mechanism in this case. Well over half of the respondents—including librarians—said that they "would prefer their online access to our journals to be through the NTP website rather than through other online service providers."

How readers read online (as opposed to on paper) is a dynamic area of research. Doyle (1986) looked at expert-novice differences in scientific journal scanning and concluded that experts tended to be more efficient and accurate than novices. Bibliographic analysis suggested differences between what Doyle terms the "hard" and "soft" sciences. Less experienced, so-called soft scientists tended to select material containing comparatively more current references. Experienced, hard scientists seemed to seek out less heavily researched topics. Two papers by teams at Xerox Palo Alto Research Center examined readers' behavior. O'Hara and Sellen (1997, p. 337) concluded that "the ability to annotate while reading was important in enforcing an understanding of the source document." Additionally, "movement" through documents in both print and online environments was important. Adler et al. (1998) looked at reading behavior based on a diary study in a variety of work settings, one of which was medical. In this sample, reading constituted an average of 70 percent of participants' work activities. The types of reading engaged in varied from skimming to more intensive attention. Observations applicable across all participants included the persistence of paper, even where computing technology was required for at least some portion of the work; the "conjunction" of reading with writing; and "cross document use." From the variations within the group, the investigators concluded that medical personnel might benefit from easy access to up-to-date information with searching and browsing capabilities (p. 248).

PUBLISHERS

Much of the information available on publishers is descriptive, reflecting the experience of a given publisher or group of publishers or reflecting observations about publishers. Indeed, in their contribution to the *Annual Review of Information Science and Technology*, Peek and Pomerantz (1998, p. 343) suggest that one reason for the "uneven" knowledge about scholarly publishing is because "advocates," typically the editors, have little time to conduct the research on electronic scholarly publishing.

Oppenheim, Greenhalgh, and Rowland (2000) conducted an opinion survey of 187 U.K.-based scholarly journal publishers with support from the Department of Trade and Industry. The questionnaire was supplemented by interviews. Of the 187 respondents, 105 were commercial companies, 71 were learned and professional societies, and 11 were university presses. Journals dominated both STM and academic publishing activities. The

Internet was indicated as the dominant format but CD-ROM “also featured strongly.” Both Internet and paper-based products represented “the greatest opportunities.” Although Internet-based activities generally were not reported to be profitable, the Internet was believed to perform one or more of three possible functions for scholarly publishing: publicity, information delivery, and (e-) commerce. Among the smaller companies (those with less than 50 employees), 83 percent considered publicity to be an important reason for using the Internet. Larger companies pointed to the importance of brand identification as an advantage in the online environment.

Questionnaire respondents maintained that the Internet would provoke a new product, enhanced by 3-D images, animation, and integrated sources and services that added to the academic’s skill set at the desktop. Opinion was mixed as to whether other kinds of information providers, such as preprint servers like that of Los Alamos National Laboratory, would reduce the importance of publishers. Over 70 percent agreed that copyright was a key concern. Over 40 percent disagreed with the statement that customers were willing to pay a realistic price for electronic products. Slightly more than half believed that their senior management failed to understand electronic publishing.

The authors point out that numerous mergers have resulted in the creation of large international conglomerates, which are believed to dominate the market. They maintain that the distinction between the commercial and not-for-profit press “is becoming less clear” (p. 362). In the survey, very large companies (those over 250 employees) were distinguished by having clearly defined electronic publishing strategies. One interviewee complained that high prices resulted from the power of the commercial publishers. When it was pointed out that the not-for-profit sector also charged high prices, the subject agreed and observed that “there had been little incentive to show restraint in a price-led market” (p. 385). This situation would change, it was felt, under competition from the Internet. (The literature on pricing is discussed later in this section.)

Kling and McKim (1999, 2000) find that publishers’ policies concerning prior publication, which may include forms of author self-posting to the Web, either to individual websites or to databases, have been a source of strain and have contributed to a movement for reform. Harter and Park (2000) conducted a questionnaire survey of 202 scholarly journals in the natural sciences, social societies, and arts and humanities in the summer

and fall of 1997 (with a response rate of 57.4 percent). They found that most journal editors do not have a formal policy regarding evaluation of work previously published in electronic form, nor are they developing such a policy. The authors found a surprisingly wide variation in policies and practice, including weight given to prior electronic publication in any form, including contribution to an online journal or set of conference proceedings. They conclude that the Internet and World Wide Web are clearly affecting manuscript consideration policies of scholarly journals.

For their part, publishers maintain that the very process of publishing adds value to the product. Morris (1999) argues that, in addition to traditional printing and production functions, publishers provide understanding of the market (important even in scholarly publishing), selection, editing, quality control, an identifiable “envelope,” branding, and added value through aids to information retrieval and navigation. Tagler (1996a, 1996b) describes efforts at Elsevier to use new information technologies to provide additional functionality in parallel with print offerings. McKay (1996) makes a similar argument, contending that quality control and organization will be the “flow-controls of the system.” Valauskas (1997) adds that “electronic scholarly journals are, like their print relatives, decidedly not about communication per se, but about validation and acceptance, so that a given idea expressed in a paper is legitimized by its publication.” On the other hand, Rous (1999), in explaining the motivation behind ACM’s decision to go online, states that there was a “perception that something is fundamentally wrong with the traditional scientific publishing process as a whole: lack of timeliness, e-print servers, more subdisciplines and specialty publications, rising prices and stressed library budgets, information overload.” He continues, “The system of recognition, promotion, and research funding is the peculiar engine for scholarly publishing. Submissions to journals increase while readership decline[s], tending dangerously toward the write-only journal.”

LIBRARIES

Tenopir and King (2000, p. 97) describe the role of libraries in the scholarly communication system as intermediaries in that they acquire journals for use by their patrons and serve as an archive. Okerson (1996, p. 199) has argued that the library is the “indispensable mediator in the dialogue between writer and reader” and will continue to have an important role; this is linked to the proposal that “universities should reclaim some responsibility for disseminating the results of faculty scholarship” (p. 198).

Given escalating costs, many libraries have looked to electronic journals as substitutions for print. While Brown (1999) found that faculty in astronomy, mathematics, physics, and chemistry seemed to want journals in print as well as online, data collected by the Max Planck Society found a clear preference for retaining electronic journals if a choice had to be made (Rusch-Feja and Siebeky 1999a, 1999b). The researchers at the Max Planck Society were also unwilling to eliminate the print versions of journals but seemed somewhat more willing to give up bound journals, suggesting that the electronic versions would constitute the archive. Goodman (2000) looked at the use of bound and unbound print journals in an academic biology library and found that parallel print/electronic formats were rarely necessary and that scientific articles might be more appropriately published only in electronic form.

Libraries have begun to question their role as intermediaries in the online environment, given the ability of their patrons to access the material directly from the publishers or from other scholars (see, for example, Jenda 1994; Goldfinch 2000, p. 244). However, the results of the Max Planck Society study as well as those reported by McKnight and Price (1999, p. 573), Pullinger (1999, p. 166–67) and Brown (1999, p. 937) suggest that librarians have an important role in archiving the published record and in educating researchers in the use of the technology as well as in identifying and raising awareness of the availability of relevant electronic journals and other resources. On the other hand, reporting on some of the results of the eLib SuperJournal project, which considered the interactions of authors, readers, publishers, and librarians, Pullinger (1999, p. 165) finds that scientists avoided the library, since it took time away from research and was generally considered inconvenient. In the summary of the project's final results, the investigators reported that use of the electronic journals in SuperJournal did result in a reduction in the numbers of visits to the libraries, but that the librarians played a critical role in the project by raising awareness of the service. Moreover, users did not view access to collections of e-journals as a replacement for libraries; rather, they saw the library "as a place to visit and browse through journals, the location of the journal archive, and where they can find helpful staff" (*Summary of SuperJournal Findings: Readers* 1999). (Results from the SuperJournal project are described more fully in a sidebar at the conclusion of this section.)

Much of the published literature covered by this literature review has appeared in library and information

science journals, evidencing the importance that this community places on these issues (see sidebar on Sample Digital Library/Electronic Publishing Projects). Many of the concerns of libraries are discussed elsewhere in this section, in the portions devoted to pricing, intellectual property, and archiving; good overviews of the issues are available in Arms (2000) and Brown and Duda (1996).

Sample Digital Library/Electronic Publishing Projects

Several early digital library/journal projects were conducted. These include CORE, which dealt with journals in chemistry (Entlich et al. 1996); Red Sage, which dealt with journals in medicine (Butter 1994, Lucier and Brantley 1995); Project Muse, which started out primarily in the humanities but includes core journals in history and the social sciences (Ekman 1996); JSTOR, which dealt with history and economics (Ekman 1996), and HighWire Press (Pudewell 1999), which describes itself as "one of the two largest free full-text science archives on Earth." Muse, co-sponsored by the Johns Hopkins University Press, and JSTOR were begun as Mellon Foundation efforts (Ekman 1996); HighWire is a project of Stanford University. JSTOR and HighWire provide conversion and archiving services to publishers suggesting ways in which new services may arise within the traditional library rubric.

A more recent effort to encourage interchange of information among collections is the Open Archive Initiative (<http://www.openarchives.org/>); this is sponsored by the Digital Library Federation and the Coalition for Networked Information, with support for technical work from the National Science Foundation and the Defense Advanced Research Projects Agency.

ISSUES IN THE LITERATURE: THE NEW ARTIFACT

Numerous studies reflect dissatisfaction with the print journal article and the need for an electronic artifact to possess new features appropriate to the domain and to the research community (see, for example, Heller 1996 and Raney 1998). In this regard, Tenopir and King (2000, p. 349), note that "The technology which makes electronic versions more than afterthought publications is the World Wide Web... We contend that in the context of scientific scholarly publishing, the Web is much more than a distribution medium because it incorporates two key

electronic elements which have the potential to revolutionize the scholarly communication system: (1) the use of multimedia applications, and (2) interactivity between authors and readers.”

In a general sense, Berghel (1999) argues that the relevant features of a new e-journal artifact fall into the general categories of enhancing contents, increasing interactivity, better access, and quality ranking and recommending systems. Whalley et al. (1996) argue that electronic journals are more flexible and can adapt to change; Hedlund, Nechitailenko, and Sears (1998) note that creating journals in electronic form with a print version as a variant of the electronic achieves efficiencies and enables elements not possible in conventional printed journals. Specific features that the e-journal article can encompass include animation and virtual reality; use of color; support for mathematical and chemical notation; hyperlinks (i.e., the links in an online document that lead either to another site or to another point in the same document) to other articles and to supporting evidence,²⁶ including algorithms, mathematical and computer routines that the user can run; visualization; multimedia and interactive displays; incremental publication, i.e., as soon as the article is ready; user-defined collections based on individual articles rather than on journals; updating; access to search mechanisms; and facilities for enhanced discourse among readers and authors.²⁷

Of the various enhancements proposed, hyperlinks have become the most ubiquitous. Boyce (1999, p. 189), who was involved in the development of electronic publishing at the American Astronomical Society, states that “an electronic journal must have copious links, for two purposes: for navigation within the article and the issue and for connectivity to other relevant resources.” Atkins

(1999) describes linking projects at the ISI Web of Science, a major indexing and abstracting service that has internal and external linking mechanisms.

Hitchcock and his teams of collaborators have investigated the use and importance of links as a means for solving some of the limitations in conventional journals. Principal among them is the need for readers, library users, and librarians physically “to get hold of all the journal articles they need when they need them” (Hitchcock, Carr, and Hall 1997, p. 115). Within the framework of the eLib Open Journal project and in cooperation with commercial and scholarly publishers, they developed a tool called the Distributed Link Service that matches a fixed database of predefined links against the contents of an article, inserting a link whenever a fixed match occurs. They compared their approach to others and expanded it in a subsequent article (Hitchcock, Quek et al. 1998) and provided an overview of their project in *D-Lib Magazine* in December 1998 (Hitchcock, Carr et al. 1998). The original work addressed citation linking; the *D-Lib Magazine* article summarizes efforts to link at the keyword level and in PDF, which is technically more challenging than working in HTML.²⁸

Jensen (1996) addresses issues of design. Design issues associated with readability (e.g., font size) surfaced in the early projects and have been of great concern to software developers with an interest in electronic books. Jensen calls attention to certain fundamental concerns, such as audience, function, resources of both the publisher and the reader, and content. He also offers specific recommendations such as the utility and placement of abstracts, tables of contents, and indices; graphics; multiple distribution modes (display, printing, ordering, etc.); and “reader specific presentation.”

ISSUES IN THE LITERATURE: PRICING AND BUSINESS MODELS

The literature on electronic journal pricing is challenging because it embraces three concurrent trends. First, there is a body of work that represents professional analysis and opinion on the appropriate balance among the interested stakeholders, in which advocates for reform and those concerned about escalating STM prices

²⁶One early study (Hickey and Noreault 1992) found a surprising “lack of enthusiasm for publishing original data with articles.” This study consisted of focus group interviews in the context of interface testing, which took place in 1990–92. The work surrounded development of an early peer-reviewed online journal, *The Online Journal of Current Clinical Trials (OJCCT)*. A description of the development of OJCCT, a joint venture between OCLC and the American Association for the Advancement of Science is provided by Brahmi and Kaneshiro (1993).

²⁷Numerous authors propound on these various elements, including Ackerman and Simonaitis (1997); Atkins (1999); Bigman and Peter (1998); Boyce (1999); Hahn (1998); Haynes (1999); Hedlund, Nechitailenko, and Sears (1998); Hildyard and Whitaker (1996); Hitchcock, Carr, and Hall (1997); Holoviak and Seitter (1997); Ihlenfeldt and Engel (1998); James (1995); Moret (1997); Steinberger (1996); Weintraub (1999); Whalley et al. (1996); Wheary and Schutz 91997); Whitaker and Rzepa (1995); and Wicks (1998).

²⁸Since the writing of this report, an article on reference linking has been published. See Herbert Van de Sompel and Oren Beit-Arie, 2001, “Open Linking in the Scholarly Information Environment Using the OpenURL Framework,” *D-Lib Magazine*, 7 (March).

have been particularly vocal. This body of work also includes economic analysis and development of the publication and business model. A second body of work represents the practitioners; this work reports on initiatives undertaken primarily by learned societies and commercial publishers and represents a body of experience. Finally, there is a small but growing body of formal econometric research covering pricing of electronic serials. (See sidebar on Pricing Microeconomics.)

McCabe (1998, 2000) provides a good introduction to the economic issues and complexities of publishing. Another good reference is several of the papers created for the March 23–24, 2000, conference on “The Economics and Usage of Digital Library Collections,” organized by Lougee and MacKie-Mason; these address the pricing of e-journals (see <http://www.si.umich.edu/PEAK-2000/>). Tenopir and King (2000, p. 236) note that “there is insufficient hard data to make any long-term predictions” concerning the dynamics of journal economics and pricing. On the more general topic of electronic publishing, Peek and Pomerantz (1998, p. 345) caution that “there is a limited base of research on electronic journals from which to draw.” With respect to studying the publishing industry (regardless of whether the product is in print or digital form), Noll and Steinmueller (1992, p. 32) point out that calculation of publisher profit margins is problematic because it requires “detailed, usually proprietary, information about each journal’s revenues and costs.” On the other hand, some consistent information on research library budgets is available (Blixrud and Jewell 1998), and the American Library Association publishes serials pricing information annually.²⁹ Thus, there is currently a fundamental asymmetry in the available information, although Bergstrom (2001) offers consistent comparisons for print journals in economics based on subscription price per page.

PROFESSIONAL ANALYSIS AND OPINION

A number of studies provide overviews of the economics of scholarly publishing. They can be divided into four general groups: articles that examine factors and causes, particularly of spiraling prices; articles that speculate about the future and set forth proposals for reform; articles that elaborate on definitions and assumptions; and articles that examine or describe the underlying economic

²⁹For the most recent prices (as of this writing), see Barbara Albee and Brenda Dingley, “U.S. Periodicals Prices—2001,” *American Libraries* 32 (May 2001):72–79; and Ajaye Bloomstone and Nancy J. Chaffin, “2001 U.S. Serials Services Price Index,” *American Libraries* 32 (May 2001): 80–81.

Pricing Microeconomics

Several issues arise within the microeconomics of the pricing itself. Fundamental notions are (1) first copy costs—i.e., the costs of producing the first copy—which are said to be 70 percent of the total cost of an academic journal; (2) the fact that in the electronic world, marginal costs, or the costs of subsequent copies, are extremely low; (3) the importance of primarily research (including research institute) libraries as the principal purchasers of scholarly journals, which means that the users (both authors and readers) do not bear the cost directly; (4) the inelasticity of library demand (which means that libraries have tended not to discontinue titles when prices increase although individuals do); and (5) the extent to which scholarly journals represent monopolies. A conflating behavioral issue arises in that individuals who abandon personal subscriptions because they rely on access from the university library do not “see” the cost that has been implicitly incurred, particularly if the library believes that it must acquire highly specialized, low-demand titles in order to serve their constituencies adequately (Stoller, Christopherson, and Miranda 1996). Additionally, Varian (1998) points out that e-journal publishing is characterized by network externality effects—that is, the value of a product depends on how many others acquire or use the product.³⁰ Varian’s observation is consistent with research that has found that a critical mass of information in digital form is necessary for readers’ acceptance of e-journals (e.g., Boyce 1999, Kelly 1997, Rusch-Feja and Siebeky 1999a, 1999b; and Zhang 1998). This point is addressed in more detail in the subsection on Studies of Behaviors and Attitudes.

³⁰Aside from Varian’s observation, we found no evidence of network externality effects having been studied in the context of e-journal systems; historically, these effects have been studied with regard to communication and transportation systems in a context of theories of natural monopoly.

and business structure. Obviously, there are overlaps among these categories, and many of the articles also contain careful discussions of the background and history of journals, of electronic journals, of libraries and technology, and of trends. Additionally, several present specific proposals—vis-à-vis the preceding discussion on changing the structure of publishing—relative to pricing e-journals. These last include work by Harnad (1998), who outlines an approach based on author page charges, an approach taken up by the Institute of Physics for its online *New Journal of Physics* (Haynes 1999); Bide, Oppenheim, and Ramsden (1998), who also discuss an approach predicated on page charges; and Varian (1998).

Among the pricing models proposed, Tenopir and King (2000) rely on observation; and Fishburn, Odlyzko, and Siders (1997), McCabe (2000), and Halliday and Oppenheim (2000) employ quantitative techniques and analyses. A conflating issue remains definition of the electronic object and its relationship, if any, to a print counterpart and/or to membership in a professional society that might include access to its publications in all formats. This conflation of print and electronic formats includes the extent to which prior experience in print publishing will carry over into electronic publishing, a perspective advanced by Tenopir and King (2000), among others.

In an article for *First Monday*, as well as in two technical papers, Varian (1995, 1996a, 1997) discusses the general problem of efficiency, differential pricing, and information. He points out that information services, like telecommunications services, are characterized by increasing returns to scale, large fixed and sunk costs, and significant economies of scope. In these contexts, the standard advice of setting prices equal to marginal costs (which are either low or zero) is not economically viable. Thus, efficient pricing will vary by service and consumer, producers will wish to engage in product and service differentiation, and profit-seeking by firms will result in differential pricing (Varian 1996a).

Varian (1997) has proposed differential pricing in the form of quality discrimination or versioning, wherein the publisher provides different levels of service “to get the consumers to sort themselves into different groups according to their willingness to pay.” Differential pricing can take different forms (Varian 1996a): different units of output sold to different customers for different prices; different units of output for different prices, but purchasers of the same unit pay the same price; and the same unit of output sold to different groups of people for different prices. Varian (1997, p. 5) describes a form of price discrimination known as “bundling,” in which distinct products are sold as a package. Bakos and Brynjolfsson (1999) set forth a model of bundling of information goods based on a menu of elements targeting different market segments. A key question in the context of pricing e-journals is how do their readers behave when given various choices. This question is investigated in some detail by MacKie-Mason and his colleagues at the University of Michigan; the early results are summarized later in this subsection. Finally, Fishburn, Odlyzko, and Siders (1997) consider the issue of pricing information goods that are likely to be consumed in large quantities,

contrasting per use versus subscription pricing from a theoretical perspective. They conclude, given users’ clear preference for the latter, that “subscription pricing is likely to dominate” (p. 1).

King and Tenopir (2000) draw a distinction between information content and the vehicle (or medium) in which the content is distributed. Most of their work has been in print journals, but they argue that understanding the underlying dynamics of print offers lessons for understanding the electronic medium (Tenopir and King 1998). They note that, from the readers’ perspective, the decision is largely a question of which vehicle; the cost of processing the information is roughly the same across all forms (King and Tenopir 2000). (This assertion is challenged by those who note that the cost of the electronic text is increased by reader demand for augmented services such as hyperlinks, as well as by those who believe that the costs of the electronic version are lower than those for print.) King and Tenopir (2000) note that, from the libraries’ perspective, however, the unit costs, when measured as cost per reading, vary depending on how heavily an article or journal is read. The costs for print and electronic journals are quite similar for heavily read articles, particularly when the researchers’ time is factored into the cost analysis and the analysis of costs varies depending on the perspective (reader or library) and components (e.g., whether scholars’ time is considered in the total cost).

King and Tenopir (2000) suggest that most of these quantitative studies “take a top-down approach using some form of multiple regression analysis to explain the large variation in price among journals” (p. 288). Among the reasons given for spiraling print journal prices are production costs, particularly in the sciences where complex graphics and the use of equations and formulae increase the costs of production; the disproportionately higher costs associated with low-circulation titles, which result from “twigging” (or the creation of more specialized journals from parent publications) and increasingly specialized journals to meet the needs of very small communities of users; monopoly power among publishers; and excessive profits charged by commercial publishers. King and Tenopir call for a bottom-up approach that begins with analyses of individual and institutional demand, based on the print model, noting further the importance of understanding the nonlinear effects of low-circulation titles; this point is also made by Noll and Steinmueller (1992).

Arjoon (1999) presents another point of view, emphasizing the economic inefficiencies of scholarly publishing and calling attention to the influence of forces that do not respond to pure market interactions: institutional conservatism, infrastructure (especially library budgets), and informational asymmetries which inhibit coordination among relevant stakeholders.

Scupola (1999) provides yet another viewpoint, examining publishing from the perspective of e-commerce in which traditional publishing is interpreted as a value chain of activities. Peters (1998) had previously set forth the value-chain approach, arguing that the sequence should be viewed in terms of functions (creation, production, distribution, protection, acquisition, organization, preservation, and utilization) rather than in terms of actors or agents (creators, sellers, intermediaries, buyers, and users). Scupola nonetheless adopts the latter framework in a model based on “complementarity,” i.e., “mutual relationships and dependence among activities whose exploration can lead to higher profitability” (p. 135). Her focus is on business processes where the information and communication technologies can be applied to production and distribution; there is no discussion of reform of publishing from the perspective of scientific communication or the values of scholarship and research.

Finally, there is a powerful argument made from the perspective of public goods. As described by Berry (2000), this is scientific research supported by funds from government and not-for-profit agencies, whose value is not *diminished* by use but, because science is cumulative, *increased* by use. Berry argues that the nature of the product—research—impels open (albeit not necessarily free) access; given the cost structure, which can make highly specialized research prohibitively expensive, he further argues that dissemination of the results “must be included in the funding of the research” (p. 39).

THE PRACTITIONER LITERATURE

Journal pricing schemes, Hitchcock, Carr, and Hall (1997) observe, typically fall into three broad categories: per article fees, per journal subscriptions, and site license subscriptions to libraries of collections (or bundles) of journals. This last scheme assumes that the publisher will make the bundle—in electronic, print, or parallel formats—available to a library for a specified number of users (i.e., the site) for an array of uses (e.g., unlimited viewing, downloading, printing on a per page basis). This arrangement requires the library to determine which journals

the publisher controls will be in its portfolio, how many users will be included in the estimate, and how the system will recognize the site and users. The latter typically relies on properties of the network (i.e., the IP [Internet protocol] address), which can exclude off-site users, depending on how the network is configured. The licenses are then negotiated on a case-by-case basis—which obviously creates some of the complexity observed in the literature. Hitchcock, Carr, and Hall (1997) observe that the market is evolving rapidly, and predictions are premature.

In an article for *iMP magazine*,³¹ Arms describes four electronic publishing models based on access: restricted access based on use (e.g., per page or per article); restricted access based on subscription (bulk access to a serial or group of serials); open access based on advertising; and open access based on external sponsorship. This is similar to the approach he has set forth for understanding the economic basis of digital libraries (see Arms 2000); he notes too that “almost every conceivable method is being tried” (Arms 2000, p. 101). This statement is consistent with observations by Spinella (2000) and with data reported by Rhind-Tutt (1998), who identified 53 commonly used pricing models for electronic products, resulting in thousands of combinations negotiated on an individual basis. Rhind-Tutt describes four principal tools:

- usage-based pricing, “which sets up a measure which assumes a certain value for each action and charges accordingly”—site licensing is a variant of usage-based pricing, whose advantages are that it is simple to administer and encourages users to seek, discover, and explore information (Rhind-Tutt, notes, however, that the “main point of contention is what constitutes a site”);
- discounts;
- value added, in which pricing differentiates among customers; and
- sponsored pricing (e.g., advertising, external funding).

³¹William Y. Arms. Economic Models for Open Access Publishing. *iMP: The magazine on information impacts*, February 2000. http://www.cisp.org/imp/march_2000/03_00arms.htm. *iMP* is not refereed and was edited by the author of this report; *iMP* articles have not been included in this literature review, given potential conflicts of interest.

Further ambiguity in licensing results from inconsistent use of terms. Thus Rhind-Tutt considers “subscription” and “site license” to be variants of usage-based pricing, but Fishwick, Edwards, and Blagden (1998) appear to use the terms differently. Arms, Rhind-Tutt, and Spinella all agree that there are numerous experiments in pricing under way, although Spinella comments that the practice in print will likely affect practice online: “Those journals relying on library subscriptions will most likely develop a library site-wide access model; while those relying on personal subscriptions or memberships may treat the online product as an added-value benefit of the print subscription.” Noting that the rule of thumb for circulation is 5,000, Spinella points out that a circulation of less than 5,000 “is probably library subscription driven. Many scholarly journals fall into this category, even those published by associations.”

Tenopir and King (2000) note that journal pricing is extremely complex and likely to become more so as a result of electronic delivery. They identify nine basic policies: no charge to the user (e.g., a journal supported by advertising or by a sponsor for promotional purposes); no charge with the understanding that there will be reciprocal service (e.g., interlibrary loan); bundled price (as part of a membership or registration at a conference); differential pricing (different prices for different categories of users); flat fee subscriptions that offer the purchaser unlimited usage rights within the constraints of the Copyright Act of 1976; fee based on an agreed-upon number of potential users; fee based on the number of simultaneous users; fee based on usage (e.g., number of file requests, downloads); and fee based on documents selected for viewing or delivery. Site licensing has been considered a reasonable contracting mechanism for large libraries and multijournal publishers. Licensing agreements can be constructed in many ways to recover costs of access and delivery.

The mid-1990s saw some interest in “micro-payments,” that is, payment on a per page or per article basis. This practice has not proved popular.³² Several studies describe the experiences of some of the prominent STM e-journal projects and initiatives (Bowen 1998, Robnett 1998, Stern 1999, Walker 1998). Boyce (1997b) outlines the reengineering process that took place at the American Astronomical Society in association with the University of Chicago Press, which resulted in substantial savings. An electronic-only copy could be offered at

70 to 80 percent of the cost of a paper subscription, and a paper subscription could be added for another 20 to 25 percent. Thus, the total cost of a parallel subscription need not cost more than the original paper subscription. Fisher (1996) compared electronic-only journals at MIT Press with paper journals, both of which were in computer science. Based on limited 1994 data, she concluded, “it seems that the direct costs of publishing an electronic journal are substantially below that of a print journal with comparable pages. The overhead costs, however, are much higher.” Moreover, “the disparity in the markets for electronic products and print products [electronic was roughly one-fifth the size of print] is, at this point in time, a very big obstacle to their financial viability, as is also the conservatism of the author community.” Fisher’s study was, however, a fairly early one, and there have been changes in the degree of acceptance as well as variability in the way that different scientific communities have adopted e-journals.

Lewis and Edwards (1998) describe the establishment of the electronic *Journal of Animal Science*, which was launched in October 1995. The American Society of Animal Science opted to include access to the journal as part of its membership fee. As of the writing of the article, the society was considering working with an aggregator to make the content available to libraries on a site basis.

Rous (1999) discusses the e-publishing initiative at the Association for Computing Machinery, which combined two elements: a publications program and a digital library. The service went online in 1997. The goal was to remain net revenue neutral. The business model combined institutional subscriptions to online publications, which meant that individual subscriptions might drop, as well as subscriptions to the digital library. ACM assumed that savings on the print side from external typesetting costs and expenditures for stock and postage would compensate for any increases associated with online production. This assumption was substantiated, although the value-added costs of the digital library may turn out to be higher than anticipated. Moreover, the business plan underestimated the service and support requirements for digital library users.

After 15 months, ACM gained 30,000 paying subscribers for the digital library. Net publications revenue rose, although printed subscriptions fell by more than 25 percent. “What comes as a surprise, though,” notes Rous, “were the number of people who did not drop their

³²Arms op cit.

print subscription when they added Digital Library access.” ACM did not see a loss in individual memberships as a result of institutional members and did see an increase in both student and international memberships; the association did not anticipate the level of consortium buying, which benefits poorer countries as well as libraries and smaller institutions to save money. Relatively few articles were bought individually.

Bot, Burgemeester, and Roes (1998) examine the cost of publishing an electronic journal, based on an assessment of costs carried out by PricewaterhouseCoopers. The general model looked at five elements (overhead, facilities, publications, creation of material, and user costs); it then distinguished between direct costs, which were assigned to the journal, and shared costs, where only part of the costs were assigned to the journal. The model focused on facilities and publications, acknowledging that the other categories represented costs that were contributed in the form of volunteer referees or were transferred to the users. The model was tested in the production of an actual law journal, and the authors concluded that electronic publishing was substantially cheaper. King and Tenopir (2000) have challenged the notion that the implied costs of authors, editors, and volunteer reviewers can be zeroed out, since this represents a commitment of time reallocated from other research responsibilities that may represent an opportunity cost or may be funded from an external source.

Peters (2000) describes the business model for *Sociological Research Online*. This effort took into consideration the needs of independent scholars—those not associated with a university or an institution—and the needs of researchers in developing countries. The business model calls for institutional subscriptions on an annual basis but is free to individuals accessing it from a dial-up account—that is, the nature of the subscription is predicated on the structure of the underlying communications technology (specifically, the IP address). A subscription permits access to a given network, and large institutions or campuses with multiple networks would therefore require multiple subscriptions. The model serves the needs of remote students who can access journals either through the network (which means they are free to the individual user) or through a dial-up account, which is also free to the user.

ECONOMETRIC RESEARCH: PRICING ELECTRONIC SERIALS

The research in this area embraces two issues: (1) pricing serials (the price of serials, particularly STM journals, has escalated); and (2) pricing information goods, given the high threshold/low marginal cost structure previously described.

Tenopir and King (2000) point out that concern over rising publishing costs and prices dates back to the 1960s. In this regard, Bowen 1998 reports on empirical research conducted in the 1980s by the Mellon Foundation, demonstrating that research libraries were facing escalating costs. The foundation believed that new information technologies might offer a partial solution and, since 1994, has supported a number of pilot projects to this end, addressing, among other items, scholarly communication and scholarly publishing. Since libraries are the primary consumers of scholarly journals, electronic publishing and the “crisis” in serials pricing have become linked issues for many, particularly those in the library community. According to Tenopir and King (2000), the price of scholarly journals increased from 1975 to 1995 by a factor of 2.6 when adjusted for inflation. In the aggregate, the cost (in current dollars) of journals has increased because the size of the journals (as measured in number of pages) has increased; they conclude that, in fact, the average cost per page has probably *decreased* as a result of computer-assisted production.

In the print world, size of circulation is clearly an issue (Noll and Steinmueller 1992). Tenopir and King (2000) emphasize the extent to which print journal costs are nonlinear—that is, the unit costs incurred by low-circulation titles (those with under 2,500 subscribers) are higher than for titles with higher readerships. Odlyzko (1999) also focuses on low-circulation journals of under 1,000 subscribers. Tagler (2000) emphasizes the threshold of size. In a discussion of the advantages and disadvantages of advertising in STM electronic journals, he cautions that advertising “really applies to a small percentage of journals fitting a specific profile” (p. 188). Specifically, advertising requires a minimum of 5,000 subscribers who are both decisionmakers and in a position to spend money; it tends to favor medicine and chemistry.

Tenopir and King (2000) compared costs across commercial, scholarly society, educational, and “other” publishers; they found that costs for educational publishers are substantially lower than those for commercial and society publishers (whose costs are comparable, based on their model). The reasons are unclear, but the researchers conjecture that these differences may be the result of university subsidies and volunteer labor. On the sensitive issue of pricing, they find that the size of journals is increasing at a time when circulation is decreasing, which drives up prices. (It is known in journal publishing that there is an inverse relationship between pricing and demand.) Since science “seems to be characterized by a highly skewed distribution of specialties representing very small scientific communities” and few very large communities, “one would expect most journals to be high priced and a much smaller number to be low priced” (p. 276). The authors find that commercial publishers generally charge higher prices, but caution that the reasons for this may be complex (pp. 319–20):

- Journal costs and prices are generally a function of size (number of pages) and circulation. Because of high fixed costs, any journal with a circulation of less than 2,500 (which is most journals) is going to have high unit costs and therefore high prices. Commercial publishers are disproportionately involved in the publication of low-circulation, specialized titles.
- Most U.S. journals in the sample appear to have been priced reasonably. There may be cross-subsidy effects in which a few “winners” support “marginal and losing journals.”
- It is not clear that university or scholarly society presses would pick up journals from commercial presses. Society presses, for example, may be unwilling to undertake publications outside their mission; and neither may have the financial willingness to mount new ventures.

With regard to the question of monopoly and pricing, there is occasional confusion over whether today’s dominant commercial publishers, resulting in part from a period of mergers, represent the monopolist or whether the monopoly exists at the journal level—that is, the circumstances are such that there is no perfect substitute, for example, for *American Economic Review*. From an economist’s point of view, scholarly journal publishing represents monopolistic competition, meaning that there are several somewhat different products, some of which

are close substitutes (Noll and Steinmueller 1992, Varian 1995). Thus, monopoly conditions appear to exist because of the relative prestige of certain journals in different fields and because of competition among authors to publish in these journals, although there is increasing evidence that some fields have adequate—if not perfect—substitutes.

McCabe (1998, 2000) extends the monopoly argument from the title level to the corporate organizational level. He argues that the behavior of the major European commercial publishers of scholarly and scientific journals can be understood within an antitrust framework. Based on his model, which he tests by looking at about 1,000 biomedical journals, he shows that the corporate mergers resulted in enhanced market power, contributing to inflated prices. A portfolio of journal titles can thus result in disproportionate market share. He suggests that future research consider other STM fields, the behavior of nonprofit publishers, and the entry of new journals into the market. In this regard, Noll and Steinmueller (1992) have examined the historical development of the market for scholarly journals based on a data set of 1,400 scientific journals. They conclude that price variation among scholarly journals “is variation in their circulation” (p. 37)—not monopolistic behavior by publishers. McCabe (1998) suggests that his findings can accommodate Noll and Steinmueller’s, since the portfolio, or aggregation, results in the monopoly position.

Tenopir and King (2000) suggest that there is substantial research to be done on the economics of journals, independent of technology, that takes into account both page count (or size) and circulation. The introduction of technology not only brings into question the role of the publishers but more generally asks how much of the traditional model carries over into an environment where content, production, and distribution are in digital form. In an earlier work, they argued that electronic-only journals saved reproduction and distribution costs, but that these savings were partially offset by the costs of storage, software, and labor (Tenopir and King 1998), a point consistent with observations by Varian (1998) and results reported by Rous (1999).

An early judgment on electronic publishing was that e-journals represented a cheaper form of publication; this notion has since been challenged. Some of the costs (e.g., of distribution) have in fact been displaced, not eliminated. Fuchs (1996) points out that scholars who have come to the conclusion that network access is free or nearly free overlook the investments that universities have made to

their communications and computing infrastructures precisely to encourage the scholarly communities to adopt the technologies in their research and to explore collaborations. Others point to the additional features that online readers seek, which add value—as well as cost—to electronic publication. Adoption of information technologies was also thought to increase the efficiencies of production processes, but estimates vary. Some observers (e.g., Day 1999) argue that the bulk of these savings was achieved in the 1980s. Varian (1998), however, suggests that savings of about 50 percent could be achieved but points out that there is a trade-off between the uniform style imposed by publishers and the idiosyncrasies of authors' word processing and writing habits. What readers will accept remains an open question, as discussed elsewhere in this report.

The research literature also discusses the cost structure of electronic journals, specifically what constitutes fixed, marginal, and variable costs; the extent to which e-journals require new services and hence incur costs; and the costs borne by users and the extent to which this affects demand. Spinella (2000), who bases his discussion on the experience of *Science*, notes that fixed costs are not uniform across all print journals (a point echoed by King and Tenopir 2000) and may include costs for marketing as well as for administrative and financial support. Archiving, in his opinion, becomes a new cost borne by the publisher (rather than the library) and is linked to quality control. Finally, there is the question of margin, or profit or surplus, which is necessary to ensure financial stability, investment capital for new projects, and funds to cover maintenance and upgrading. Spinella concludes that “endless arguments may ensue, about how much margin (profit) is ‘permissible’ in a scholarly journal, or even whether any margin should be charged by nonprofit entities. In the end, it will be acknowledged that nearly all important and vibrant publications will charge some sort of margin.”

Halliday and Oppenheim (2000) discuss results of three models for electronic journals from a U.K. perspective: traditional (similar to print); noncommercial, based on work by Harnad who has posited an author-paid model; and a free market model, based on work by Fishwick, Edwards, and Blagden (1998). The third model called for a combination of author payments and sales and external foundation sponsorship to cover costs for editors and reviewers; there would also be a system of royalties for authors. The intent of the model was to introduce feedback mechanisms and incentives that would create an efficient market for scholarly articles while

reducing prices. As a result of running the simulation, Halliday and Oppenheim concluded that traditional journals on the print model (where there was substantial contributed labor) could be run quite cheaply with a minimum of 500 subscribers. The author-supported model was also viable, but the authors believe that the barrier to implementation is cultural: author fees are not popular outside of the United States. The free market model was shown to be price competitive with the traditional model, but the authors have chosen not to pursue it. Future work calls for examination of a “cottage industry” model—i.e., journal production in a modest organization—and model publishing in large organizations to examine the role of organization size with respect to cost.

The PEAK project, housed at the University of Michigan, is both a production service for electronic journal delivery and an opportunity to conduct pricing research based on approximately 1,200 journals published by Elsevier Science, which were provided in full text to 340,000 users through libraries at 12 campuses and commercial research facilities. The project ran from January 1995 to August 1999, during which time the investigators looked at bundling and pricing and how the economic behavior of users of electronic journals might affect pricing strategies and libraries' decisions. The PEAK website provides a complete description of the project (<http://www.lib.umich.edu/libhome/peak/>); several papers and articles also describe the experiment, preliminary results, and theoretical underpinnings (Bonn et al. 1999, Gazzale and MacKie-Mason 2000, MacKie-Mason and Jankovich 1997, MacKie-Mason et al. 1999, MacKie-Mason and Riveros 1997).

The design offered users packages containing two or more of three “access products”:

- traditional subscription: unlimited access to the material also available in print;
- generalized subscription: unlimited access to any 120 articles from the entire database, selected on the basis of user requests; and
- per article: unlimited access for a single individual to a specific article.

As of this writing, analysis of the PEAK project is under way. Among the preliminary conclusions was confirmation that the generalized subscription model is only feasible in the electronic environment but is quite successful; the user cost of access, whether monetary

payment or convenience, has a substantial effect on the number of articles that readers access; there is a substantial learning curve; and recency is important. The researchers have further concluded that electronic publishing can expand access through innovative approaches such as the generalized subscription, which, at the same time, enables publishers to obtain a predictable revenue flow (MacKie-Mason et al. 1999, Gazzale and MacKie-Mason 2000).

ISSUES IN THE LITERATURE: INTELLECTUAL PROPERTY RIGHTS

Maskus (2000) defines intellectual property rights as ownership and control of an idea or piece of information that may have economic value if put to use in the marketplace (p. 27). In the United States, intellectual property rights are protected in the Constitution (article 1, clause 8) as well as by Federal legislation (Copyright Act of 1976, Digital Millennium Copyright Act of 1998 [DMCA]). The protection of intellectual property rights in an analog environment is predicated on a tenuous balance between the communal values of scientific research and the exclusionary values of ownership and is based on a definition of “copy”—that is, what entity, under what circumstances and conditions and for what purposes may make copies of a given work and reap the economic benefits of making those copies. Intended to be flexible, the law accommodates a range of behaviors and actions (Burk 2000, Day 1999). New information technologies, however, with their capacity to diffuse information goods rapidly, enables making copies and making changes to those copies and thus upsets this balance by dramatically simplifying mass infringement (Burk 1993, Elkin-Koren 1995, Lai 1999, Soon 2000, and Zamparelli 1997). As Kelsey and Schneier (1999) summarize, “the

fundamental enforcement problem is that...nearly anyone with a computer and an Internet connection will be capable of posting copyrighted materials to the Internet [which] once posted, will be retrievable by nearly anyone.”

While some authors have characterized technological developments as a “crisis,” Arms (2000) notes that similar concern over the interaction between economic issues and copyright law has occurred with the introduction of every new technology. (For a discussion of the tenor of the debate, see Berry 2000 and Smaglik 1999.) Intellectual property rights continue to be actively discussed in legal, policy, and economic circles as well as among researchers. In 1997, the Computer Science and Telecommunications Board (CSTB), with support from the National Science Foundation, conducted a study of intellectual property in the information age; the resulting report, *The Digital Dilemma*, is an excellent introduction to the issues (CSTB 2000). Pamela Samuelson’s “Selected Tutorials for Non-Legal Professionals and Other Slide Presentations” (see <http://www.sims.berkeley.edu/~pam/tutorials.htm>) is another helpful introduction. Useful websites for understanding and monitoring developments in this complex and fast-changing area of the law are listed in table 1.

The literature examined in this report divides this most recent iteration of the technology and intellectual property debate into two general, but interrelated, categories: the law itself, and whether and how new technology challenges the law’s assumptions and potentially enforces and/or obstructs it. Legal argument has its own conventions, and although the legal literature itself was not covered by this study (in the sense of examining case law, for example), it is clear that writers reflect the traditions of

Table 1. Web resources for intellectual property rights

Name of site	URL
FindLaw’s Intellectual Property Law	http://www.findlaw.com/
Cornell University Legal Information Institute	http://fatty.law.cornell.edu/
Copyright Resources Online, by Ann Okerson	http://www.library.yale.edu/~okerson/copyproj.html
Stanford Technology Law Review	http://stlr.stanford.edu/STLR/Core_Page/index.htm
World Intellectual Property Rights Organization	http://www.wipo.org/
The American Bar Association’s Section of Intellectual Property Law	http://www.abanet.org/intelprop/home.html
American Intellectual Property Law Association website	http://www.aipla.org/
Library of Congress	http://lcweb.loc.gov/copyright/
Berkeley Digital Library SunSITE: Copyright, Intellectual Property Rights, and Licensing Issues	http://sunsite.berkeley.edu/Copyright/
iKnight Technologies’ Intellectual Property Law Server	http://www.intelproplaw.com/

legal analysis (reliance on precedent, projecting outcomes based on the logic of an argument, etc.).

Agreement in the literature stems more from definitions of the problem(s) than probable solutions. Burk (2000) identifies a possible cause for this uncertainty, stating that the “rapidity of change for both the relevant law and technology” makes any analyses of the subject rather transient (p. 16). Other authors (Kleinman 1996, Okerson 1991) agree with Maskus’s (2000) contention that Western notions of intellectual property are drawn from three broad philosophies on the nature of intellectual property and its protection. The *natural rights* view (also called *moral rights*) is a European tradition that “assigns ownership of mental creations to their inventors” (p. 27).³³ The *public rights* view prohibits the assignment of private property rights to intellectual creations, arguing that “free access to information is central to social cohesion” (p. 28). The *utilitarian* view, adopted by many countries (including the United States), recognizes that intellectual property rights may be assigned and regulated for purposes of social and economic policy and aims to balance the benefits and costs of property rights in information (p. 27). The tensions among these views inform some of the debate over business and economic models (discussed above).

Burk (2000) parses specific intellectual property protections into five major modes, three of which are common across the literature examined in this study: patents, copyrights, and trade secrets. Trade secrets, like other forms of proprietary information, are most relevant to scientists who work in corporate laboratories and can affect their willingness or ability to publish (see also Walsh and Bayma 1997).³⁴ While each of the different modes of intellectual property protection applies to various aspects of electronic scientific journal publishing, patents and copyright appear to be the most contentious areas. For a comprehensive bibliography on electronic patent and copyright issues, see <http://cyberlaw.stanford.edu/lessig/>.

PATENTS

Patents are defined as “exclusive rights granted by the federal government to the inventors of new and useful machines, articles, substances or processes” (Burk 2000, p. 19). Disagreement over patents in electronic scientific scholarly publishing may arise with regard to

the use of patented software, hardware, or processes as part of the electronic dissemination, “fencing off,” or sale of scholarly materials (p. 19). Patent law is especially controversial in the context of electronic scholarly publishing in such fields as biotechnology, where published works often have immediate and lucrative market applications (Eisenberg 1987, Walsh and Bayma 1997). In particular, traditional publishing practices reflective of communal scientific research norms (e.g., requirements to prerelease, publish, and/or provide access to underlying data) have been reexamined and, in some cases, revised to reflect this new commercial reality (Eisenberg 1987, Walsh and Bayma 1997). Although lauded by some, others view this strengthening of “ownership” as detrimental to scientific research “writ large” (Burk 2000, p. 27).

A separate, but potentially important, debate questions the wisdom of patenting software and so-called business practices such as online auctions (e.g., eBay or Priceline), which often incorporate common computing and networking processes (Burk 2000). Related to business practices are issues that arise in the context of the business practices and contracting, which is largely governed by the Uniform Commercial Code (UCC) in the United States. In terms of software, producers and publishers of various content have sought changes to the UCC through the controversial Uniform Computer Information Transactions Act (UCITA) as a means to enforce so-called shrinkwrap licenses.³⁵ Shrinkwrap, or more recently “clickwrap,” licenses assume that “the consumer, by opening the packaging of the software,” or clicking an “I agree” on-screen icon, “has shown an intent to be bound by the terms of the license governing the software” (as cited by Burk 2000).³⁶ The legality of this approach has yet to be settled in Federal court, but

³⁵The controversies surrounding development and enforcement of UCITA by the states may be followed at websites maintained by the Association for Computing Machinery (<http://www.acm.org/usacm/copyright/#ucita>).

³⁶Shrinkwrap and clickwrap agreements are used to avoid the “First Sale” doctrine in 17 U.S.C. 109, where a buyer of software normally has the right to resell, freely exchange, or lease or loan software. These agreements seek to control a buyer’s use of software by licensing specific rights and expressly prohibiting reverse engineering. Shrinkwrap licenses are disfavored due to their condition-subsequent nature, and, although originally held to be generally unenforceable (see *Vault Corp. v. Quaid Software Ltd.*, 655 F. Supp. 750 [E.D. La. 1987] aff’d, 847 F.2d 255, 268–70 [5th Cir. 1988]), recent trends have favored enforceability (see *ProCD, Inc. v. Zeidenber*, 86 F.3d 1447 [7th Cir. 1996]), holding that shrinkwrap licenses are enforceable contracts. See also *Hill v. Gateway 2000, Inc.*, 105 F.3d 1147 (7th Cir. 1997). Clickwrap may make shrinkwrap litigation moot, due to its noncondition-subsequent nature.

³³This tradition is not universally adopted in European countries. For example, it is not a part of the U.K. legal code.

³⁴In this context, note that, in the event that trade secrets are published, both the publisher and the author are liable.

these examples represent an early iteration of attempts at legal standardization (Burk 2000).

Regarding patenting business practices, online service providers have tried to expand notions of ownership, including one that claims ownership of hyperlinking. It is argued that the logistics of enforcing such patents potentially overburden the judicial system, especially given the already significant integration of computer processes such as hyperlinking. Moreover, many believe business practice patents will stifle e-commerce development. The outcome of these related debates could affect scientific electronic publishers interested in expanding their online presence (e.g., with interactive multimedia) beyond merely digitizing their print products.

COPYRIGHT

According to Arms (2000, p. 114), U.S. copyright applies “to almost all literary works, including textual materials, photographs, computer programs, musical scores, and video and audiotape.” Copyright involves giving up specific freedoms and retaining others (Okerson 1991, Stallman 1996). Initially, the creator of the work—or, depending on the terms and conditions of employment, the employer of the creator (with the exception of government employees)—owns the copyright for a finite term and can make use of or dispose of it (e.g., through sale or license) like any other property (Arms 2000). There exists what is called a “bundle” of rights conventionally associated with copyright; these rights may include translations, creation of derivative works, serial rights, and distribution, among others.³⁷

Day (1999) contends that copyright is equated with property rights as a means to balance the “incentive effect” and the “enabling effect,” the two principles that underpin the motivation for creating and the means for

benefiting (financially or reputationally) from the dissemination of copyrighted works. Tenopir and King (2000) note that most scientific researchers publish for reasons associated with communication, prestige, and career advancement; thus, the financial incentive has historically been less important to them, except in those more lucrative fields such as biotechnology. Okerson (1991, p. 425) observes that this unique perspective of the scholarly community might prompt a renewal in interest by universities in a less commercial distribution and ownership of ideas to facilitate the broader sharing of ideas.

First sale and fair use, the dominant concepts of copyright, distinguish between a copy and the original work and regulate how both are managed (Arms 2000, Burk 2000, Okerson 1991). According to Arms (2000, p. 118), “First sale applies to a physical object, such as a book, that the copyright owner may sell, but once the customer purchases a new book, that customer has full ownership of the copy,” and may dispose of it (including through sale) without having to get permission. “Fair use,” he continues, “is a legal right that allows certain uses of copyright information without permission of the copyright owner.” While the right to quote short passages and copy portions or articles from anthologies is permitted, the boundaries of fair use are “deliberately vague.” Arms concludes that, “in general, fair use allows reproduction of parts of a work but not the whole, single copies rather than many, and private rather than commercial use” (p.118). Disputes arise over copyright in an electronic environment because technology, by facilitating and simplifying copying, sharing, distributing, and modifying, mitigates the controlling effects of such concepts as first sale and fair use, and tests the previous consensus around notions of “some” infringement by making most infringement virtually effortless (e.g., “cut and paste” functions, which can be applied to documents found on the Web and saved by users.

Given that U.S. law is predicated on the notion of “copy” (Zamparelli 1997), another area of heated dispute relates to the status of so-called RAM copies (Hardy 1997).³⁸ Whenever digitized information is

³⁷The copyright owner has the right, pursuant to section 106 of the 1976 Copyright Act, to reproduce the copyrighted work in copies or phonorecords; to prepare derivative works based upon the copyrighted work; to distribute copies or phonorecords of the copyrighted work to the public by sale or other transfer of ownership, or by rental, lease, or lending; to perform the copyrighted work publicly, in case of literary, musical, dramatic, and choreographic works, pantomimes, and motion pictures and other audiovisual works; to display the copyrighted work publicly, in case of literary, musical, dramatic, and choreographic works, pantomimes, and pictorial, graphic, or sculptural works, including the individual images of a motion picture or other audiovisual work; and, in the case of sound recording, to perform the work publicly by means of a digital audio transmission. The copyright owner can license all or a portion of these bundle of rights. However, section 107 of the act imposes a reasonableness standard, or fair use provision, that limits the otherwise broad scope of the use of the bundle of rights.

³⁸As defined by the Computer User High-Tech Dictionary (<http://www.computeruser.com/resources/dictionary/index.html>), “random access memory,” or RAM, “is the memory used for storing data temporarily while working on it, running application programs, etc.” “Random access” refers to the fact that any area of RAM can be accessed directly and immediately, in contrast to other media such as a magnetic tape where the tape must be wound to the point where the data (e.g., a song) are located. RAM is called volatile memory; information in RAM will disappear if the power is switched off before it is saved to disk.

accessed in automated information processing and retrieval systems, it is reproduced and stored in computer memory—most often in RAM or in temporary “cache” files on magnetic media. Temporary copies are also made during transmission as digital documents requested by users are routed to their final destination (Burk 2000). Because these copies are accessible to system operators, for example, and may often be permanent, some courts believe their mere existence constitutes infringement of copyright. To the extent that such unauthorized copies are redistributed, the normal operation of computers and computer networks consequently may result in repeated and widespread copyright infringement (Litman 1996b).

Such reasoning has many critics, including Jaszi (1996) and Samuelson (1996). Concepts of fair use and implied license are applicable as potential solutions. Fair use, which permits access if it occurs in a manner consistent with the author’s intentions, might view placing information on an open network as indicative of communal intent. Similarly, implied licenses—presumed license arrangements that are inferred from the actions of the party authorized to make such arrangements, or an inverse of the clickwrap license mentioned above—might also view placing material on an open network as implied permission to generate and distribute RAM copies (Burk 2000).

Finally, there is wide disagreement over copyright associated with hyperlinking (Burk 2000). The discussion turns on whether the URL—i.e., the string—itself is subject to copyright and whether the document that is retrieved for the user as a result of selecting the copyrighted string is, in some way, incorporated into the original document and whether that process constitutes an infringement of the rights of the owner of the second document (see, for example, Cavazos and Miles 1997). Thus, one way of understanding the challenge posed by hyperlinks is to see the process from several perspectives: that of the reader (satisfaction), that of the author (prestige), and that of the publisher (profit).

Arms (2000), Bennett (1999), and Burk (2000) are among the many who have observed that most of the literature focuses on the norms and laws of the United States, rather than of other countries, regarding intellectual property. Burk explains that this might be due to this country’s “significant lead in the development of legal precedent related to digital networks” (p. 16). Our study, however, uncovered some notable research on other countries. Cheverie and Trump (1997) note that the European

Union maintains a number of comprehensive websites for its copyright and intellectual property-related research projects (see especially the sites for the Telematics for Libraries Program, <http://www2.echo.lu/libraries/en/libraries.html>; and the European Copyright User Platform <http://www.kaapeli.fi/ebliida/ecup/related/index.html>). Luzi (1998) discusses some non-U.S. copyright and peer review norms in the context of “grey” literature in Europe and the implications of conflicting legal and cultural jurisdictions. Arms (2000), in an overview of legal issues for digital libraries, discusses how overlapping national and international jurisdictions increases the number of applicable laws and complicates enforcement of cross-national norms regarding such issues as free speech and liability. Lai (1999) and the World Intellectual Property Organization (1999) discuss similar issues in the context of the impact of noncircumvention clauses on global networked services and the enforcement of the Trade Related Aspects of Intellectual Property Rights Agreement, respectively. Burk (1993) examines how pre-Web global networks complicate enforcement of software patents, and Maskus (2000) discusses the need to examine non-Western intellectual property norms and laws in the context of furthering trade relations.³⁹

POSSIBLE TECHNOLOGICAL SOLUTIONS

In general, technical protection mechanisms are considered “useful” in protecting intellectual property rights, but not a “panacea” (CSTB 2000, p. 13). How technology reinforces ownership of information is frequently discussed as “access management” (Arms 2000), “access control” (Gladney 1997), or technical protection services (CSTB 2000). Overviews of the technological issues can be found in Arms 2000, CSTB 2000, Gladney 1997, Gladney and Lotspiech 1998, and Stefik 1999; the last defines the basic issues and posits a technological solution (chapters 2 and 3).

Despite criticism by Kelsey and Schneier (1999) that deference to technological solutions for copyright enforcement invites censorship, access management provides copyright owners with some practical options for

³⁹Although it is outside the immediate concerns of scholarly publishing, a number of authors noted that recent discussions of privacy and the provisions of safe harbor illustrate both the international scope of these issues as well as the ability of nations to arrive at a process for resolution. (For detailed information, see the U.S. Department of Commerce safe harbor website at <http://www.export.gov/safeharbor/>.) Recent cases in France concerning the sale of World War II era memorabilia and what is considered “protected speech” in the United States also highlight tensions among nations over acceptable public behavior.

managing their information “property.” In general, Stefik (1999, p. 55) argues, access management of commercially valuable content requires some type of “trusted system,” or, in Gladney and Lotspiech’s (1998) term, a “trustworthy system”—a system programmed with machine-interpretable digital rights language that protects and governs the use of digital objects and information. Examples include Xerox Corporation’s Digital Property Rights Language and IBM’s Document Access Control Method. Such systems, according to Stefik, facilitate a “microtransactions” approach to fair use.

Trusted systems are designed to prevent large-scale unauthorized copying of protected works. CSTB (2000) summarizes some of the necessary components for effective access management—the combinations of which are many—as follows (p. 155):

- *Security and integrity features of computer operating systems* include, for example, the traditional file access privileges enforced by the system. These can also include measures to ensure message and transaction authentication, integrity, and nonrepudiation.
- *Rights management languages* express in machine-readable form the rights and responsibilities of owners, distributors, and users, enabling the computer to determine whether requested actions fall within a permitted range. Charging mechanisms can also be part of rights management rules. In general, these languages can be viewed as an elaboration of those used to express file access privileges in operating systems.
- *Encryption* allows digital works to be scrambled for subsequent unscrambling by legitimate users only.
- *Persistent encryption* allows the consumer to use information while the system maintains it in an encrypted form.
- *Watermarking* embeds information (e.g., about ownership) into a digital work in much the same way that paper can carry a watermark. A digital watermark is sometimes called a “tracing object” because it can help owners track copying and distribution of digital works.

Arms (2000) and Gladney and Lotspiech (1998) describe one variation of how these elements are combined in trusted systems, IBM’s Cryptolope system, which uses secure “containers” to transmit information across the Internet. Information is transmitted in a cryptographic envelope called a container, in which information suppliers seal their information. Recipients can open the container only after they have satisfied any access management requirements, such as paying for information use. Although the envelope can be passed on to others, they too must pay to obtain the code to open the envelope. Cryptolopes sometimes contain an abstract in clear text that provides users with a description of the content and the terms to which users are expected to adhere if they view the content. Arms notes that the lack of widespread deployment of a public key infrastructure presents a significant barrier to the acceptance of the Cryptolope system.

The success of a technical protection service begins with its inherent technical strength but also depends on the product it protects and the business in which it is deployed. According to CSTB (2000), factors such as system usability and the appropriateness of the system to the content and the threat are the properties that bring a technical protection service in line with a business model.

Overall, use of trusted systems to enforce terms and conditions provides a much finer grain of control than the law and moves the legal basis of protection toward that of contracts and licenses (Stefik 1999). Samuelson (1996) expands on this idea by suggesting that creating consumer-friendly business models based on technological solutions, rather than reinventing or strengthening copyright law, is the “biggest challenge that cyberspace poses for authors and publishers.” Two major components in developing such information marketplace models are to design systems that can visibly demonstrate their integrity and accountability and to create institutions that can certify these trusted systems, perhaps along the banking model of an accounting paper trail that backs up the occasional failure (Stefik 1999).

The Digital Millennium Copyright Act (DMCA) of 1998 addressed many of these concerns about ownership of intellectual property rights, use of intellectual property, and protection from theft, including liability of online service providers, libraries, and educational establishments; copies of materials made by service providers for technical reasons (e.g., caching or retransmit); and the

circumvention of technical methods used by copyright owners to restrict access to works (Arms 2000). The provisions of the DMCA were not uncontroversial and resulted from significant debate within and among the government, legal, and computer science communities that began in the early 1990s with the Clinton Administration's Working Group on Intellectual Property Rights, part of the Information Infrastructure Task Force. An initial "Green Paper" and a subsequent final "White Paper" were produced by the group and served as a partial basis for the DMCA legislation. For an overview of the complex provisions contained in, and the debate surrounding, the DMCA, see the special 1996 issue of the *Oregon Law Review*, "Innovation and the Information Environment." While most of the articles criticized the final White Paper (Jaszi 1996, Litman 1996b, Stallman 1996, Yen 1996), one (Glisson 1996) argued that the White Paper's recommended changes were not, in fact, dramatic. Additional criticism of both the Green and White Papers came from Gassaway (1996) and Samuelson (1996). Elkin-Koren (1995) provided criticism of the Green Paper, noting that the new technology created an "opportunity for social change" and the decentralization of information flows (p. 196).

Two case studies that precede the DMCA but illustrate the sometimes counterintuitive complexity of copyright enforcement are Schweighofer (1997) and Conley and Bemelmans (1997), which examine the implications of copyright law for a legal firm's use of research databases and the creation of a fine arts educational CD-ROM by a museum, respectively. Sully (1997) describes her experience with intellectual property rights enforcement on the JSTOR project.

As of this writing, the computer science community maintains that the DMCA's prohibition on circumventing technical means of enforcing digital copyright thwarts information security innovation. For differing stakeholder perspectives on this aspect of the debate, see the ACM Intellectual Property Library (<http://www.acm.org/usacm/copyright/>), the Electronic Frontier Foundation Digital Millennium Copyright Act Archive (http://www.eff.org/pub/Intellectual_property/DMCA/), the ARL Digital Millennium Copyright Act Status & Analysis website (<http://www.arl.org/info/frn/copy/dmca.html>), and the American Library Association's Copyright and Intellectual Property website (<http://www.ala.org/washoff/property.html>). The key to resolving these tensions, according to many authors, will be how the courts interpret the law.

Ultimately, at the heart of the intellectual property rights debate is the difficulty of balancing and/or reconciling the interests of very diverse stakeholders in digital publishing under the jurisdiction of one copyright or intellectual property regime (CSTB 2000, Samuelson 1996, Stefik 1999). The interests, conventions, and practices of groups such as musicians and filmmakers can differ dramatically from those of print publishers and academics.⁴⁰ Changes in the technology have rendered previous distinctions regarding an unnetworked "artifact" or "display form" (print versus audio recording versus film) less meaningful—and therefore less helpful in enforcing various interpretations of first sale and fair use. Given the fluidity of environments on a number of related fronts (technological, user, and industry), most authors agree with the CSTB (2000, p. 239) conclusion that "major change in public policy and regimes at this stage is ill-advised."

ISSUES IN THE LITERATURE: PEER REVIEW

Peer review, which some observers posit as the boundary between formal and informal scientific communication (Sandewall 1997, Schauder 1994, van Raan 1997), is a frequently discussed issue in the literature. The American Association for the Advancement of Science (1998) notes the importance of peer review and adequate quality control to electronic publication,⁴¹ given "the ease of publication"; this point is echoed by Hahn (1998) and was also made by ICSU Press/UNESCO (1996).

Roberts (1999) argues that the volume of information available today, particularly on the Internet, means

⁴⁰Note that a body of contract law exists that affects specific conditions under which given parties agree that a work may be used and distributed. Since most academic authors sign away their copyright as a condition of publication, the issue of contract law and how contracts might be used to modify existing practice appears not to have arisen at this point within the framework of scholarly publication. However, this is very much an issue within the framework of commercial publishing, where electronic rights to works are carefully negotiated as are film rights and other future uses of a work. Whether a variant of these practices migrates to the world of scholarly publishing remains to be seen.

⁴¹Quality control in this regard implies not only attention to content issues (e.g., proofreading, ensuring consistency of format and presentation), but also to issues related to digitizing information and archiving electronic materials (e.g., addressing concerns regarding the permanence of materials, system reliability, the possibility of corruption, transmission errors, and the effects of malicious or careless behavior).

that “standard refereeing practices are...likely to remain an important mechanism for allowing readers to make meaningful distinctions between reputable scholarly work and second-rate material.” Moret (1997) makes a similar point, noting that “in mathematics, the hard sciences and engineering...peer review remains the single mark of quality—and no viable alternative has yet been proposed.” He points out, however, that the technical infrastructure for peer review is not free and must be borne by the publisher. Harnad (1999) does not agree; he argues that, while peer review is necessary for quality control and certification, robust digital communications technologies and archiving (along the model established by the Los Alamos preprint server) eliminate the need for publishers—assuming that authors are willing to give their work away; this last is a characteristic of the learned research literature, but not necessarily of the “bigger world of trade publishing.” Peer commentary, he believes, is a supplement to formal peer review (Harnad 1999). Singer (2000), discussing biomedical research, also calls for multiple forms of evaluation, including citation analysis, secondary reviews, and “new open forms” of peer review.⁴²

Other researchers go beyond this suggestion, advocating substitution of a community-based approach to peer review, as well as a decoupling of peer review from formal publication. One of the more extreme proposals is set forth by Nadasdy (1997), who organized the *Electronic Journal of Cognitive and Brain Sciences*. Nadasdy’s journal accepts articles for publication before review; to create and maintain standards, it uses a “two-tier acceptance procedure that makes reviewing automatic and allows readers to control final acceptance.” Articles with a rating of 80 percent or higher based on a short online evaluation are transferred to the archive at the end of the month; the others are removed. A variation on this idea is proposed by Stodolsky (1995), who posits a “consensus” journal with reader-submitted reviews; and Sumner and Shum (1996), who set forth a system that allows computer-supported collaborative argumentation to be built into a reviewing software system as part of the *Journal of Interactive Media in Education*.

⁴²In this regard, Schoonbaert (1998, p. 98) appears to argue for traditional peer review, and perhaps against accessible e-publications, in the biomedical sciences because of public health implications: “Medicine is not primarily focused on theory building but affects human lives in a very tangible way. Thus its clinical nature may be compromised by reaching too general a lay public, who might be especially prone to misinterpretation or unsound use of unfiltered information.”

These and other proposals to reform or replace the peer review process reflect more than just the effects of expanded e-publishing. Objections to peer review are in general independent of the technology; critics cite the conservative bias of the process (Roberts 1999) and lag. Lag, it turns out, appears to be a function not of technology but of reviewers’ timeliness (Moret 1997, Tenopir and King 2000).⁴³

The literature contains various studies of aspects of peer review and electronic publishing. Wood (1998) reports the results of a combination of interviews, a survey questionnaire mailed to 200 academic members of editorial boards (the response rate was not provided), and a focus group of five biomedical academics. Fields covered included biochemistry, endocrinology, medicine, microbiology, and immunology. The respondents used either Word or WordPerfect for word processing and a variety of software tools to produce graphics; there was also variation among their e-mail systems, which affected the ability to transfer files. Nevertheless, the authors were enthusiastic about submitting manuscripts electronically. Overall results demonstrated a widespread interest in and support for using e-mail to enable broad-based collaboration and peer review with panels assembled from an international pool. Interviews revealed an emphasis on the global nature of peer review, which permitted “authors and referees in less accessible countries a chance to contribute to the process” (p. 195). Wood and Hurst (2000) describe a system for online peer review in the biological sciences, intended in part to identify specifications that would be acceptable to users while cutting costs for publishers, improving their services, and streamlining the process. Responses (59) to the experimental system were positive.

Pedersen and Stockdale (1999) sent a questionnaire to department heads at seven universities in the United Kingdom (10 percent of the nation’s total) with strong reputations in mathematics and science, since these fields were early adopters of electronic communication, including journals. The survey used a cluster sampling methodology (the total sample size was not given) and was followed up by interviews. The researchers focused primarily on users of electronic journals; they found that “all those interviewed felt that the peer review process was essential and, although, it did slow down the publication process, worked well” (p. 49).

⁴³The July 15, 1998, issue of the *Journal of the American Medical Association* was devoted to peer review, although it did not address the implications of the information technologies explicitly. For a summary of the contents of the issue, see Rennie 1998.

Singleton (1997) describes the results of a survey in which the Institute of Physics requested information on a range of topics including respondents' roles as authors, referees, readers, influencers, and purchasers. Of the 13,000 surveys sent to physicists worldwide, 3,500 were returned. The author reports that, in the institute's experience, 30 percent of submitted papers were accepted and 60 to 70 percent of these were revised, "sometimes substantially so." One question asked respondents to rank whether "material improvement resulted in the papers you have refereed." Less than 1 percent believed that this had never occurred, and 14 percent believed that this had always occurred; the remaining responses were distributed between these two poles. Respondents were next asked, "what percent of your articles materially improved" as a result of the review process. To this question, 8 percent of the respondents replied that 0 percent improved, and 15 percent responded that 76 to 100 percent improved; the remainder were again distributed between these extremes. Other questions in the survey concerned attributes of electronic journals, equipment used, and the library journal selection and acquisition process.

Sweeney (2000) conducted a survey of the Florida State University system. He contacted 75 individuals and received 62 responses, evenly divided between senior administrators and faculty. The study results are ambiguous. Sweeney asked if the peer review process was as thorough in electronic journals as it was in print. Both administrators and faculty were approximately equally divided on the issue, although one commented that, "all e-journals are not equivalent in many ways. You really need to define very explicitly and specifically what you mean by 'electronic journal.'" When asked whether electronic publishing undermined the integrity of academic rigor, both the administrators and the faculty disagreed or strongly disagreed. Comments indicated that "this statement was contingent upon the demonstrated quality of the peer-review process associated with electronic publication." There was no overwhelming indication that issues relating to electronic publishing were more important for tenure-seeking faculty than for tenured faculty.

ISSUES IN THE LITERATURE: ARCHIVING

The relationship between publication and archiving has been observed by those concerned with electronic publishing (e.g., Ekman 1996 and Garson 1998) and contributed to the organization of JSTOR (Bowen 1998).

The relationship raises questions of permanence, reliability, and acceptance (AAAS 1998, ICSU Press/UNESCO 1996). Zhang (1998), for example, found that the fluidity of the Internet and the continued accessibility of e-sources—not just of e-journals but also of mailing lists (i.e., listservs) and newsgroups—constituted a barrier to their acceptance. Odlyzko (1996a) believes concerns regarding the durability of electronic information are "unjustified." However, he treats the problem primarily in technical terms (e.g., storage capacity of disks and stability of CD-ROMs); he does not go into the organizational and managerial implications of creating, updating, and maintaining an archive, a point emphasized by Boyce (1999). HighWire Press (<http://highwire.stanford.edu/lists/largest.dtl>), which holds extensive journal collections in biomedicine and other sciences,⁴⁴ provides archiving and other electronic services for publishers and, based on the terms reached with the individual publisher, makes the archive of a given journal free or open to users.

Sandewall (1997), Smith (2000), and Smith (1999) discuss the proposal that the electronic journal act as an "overlay" on archives or databases of preprints. As noted earlier, hyperlinks among collections—whether of articles, reports, or data—represent a consistent feature of the new electronic artifact, and the relationship between the article and the archive was an important feature in the development of electronic journals by the American Astronomical Society (Boyce 1999) and the American Physical Society (Kelly 1997). Researchers at the Max Planck Society seemed willing to give up print versions of backfiles if an online archive were maintained (Rusch-Feja and Siebeky 1999a, 1999b). However, one respondent to the survey noted, "The guarantee of maintaining a comprehensive archive is essential to any theoretical research." The U.K. SuperJournal project (see sidebar at the end of this section) also found that a backfile (or archive) of electronic journals was an important feature for users (*Summary of SuperJournal Findings: Readers* 1999).

Incremental publication—i.e., publishing when the article is ready rather than when the publisher is ready—is one argument on behalf of electronic publishing in the context of publishing reform. The ability to update research is another. These conveniences pose a significant problem from an archiving standpoint, however, with regard to what versions should be retained. As Smith

⁴⁴As of November 27, 2001, 339,813 articles in electronic form were available free of charge through HighWire Press.

(2000) has argued, in sciences where progress is incremental and investigators build on generations of results, a stable record of the succession of results—including those challenged in subsequent experiments—is an important dimension of the conduct of research. Thus, there is an inherent tension between the need to reflect currency and revision and the need to maintain the record; this conflict, Smith notes, may be resolvable through a combination of reliable archiving and linkages. Eason et al. (1997) qualify the point; they find that “the cumulative character of the hard sciences [e.g., chemistry, molecular genetics and proteins, physics, geography, engineering] means that the dominant concern is with recent knowledge whereas the re-interpretative character of the soft sciences [e.g., history, archaeology, psychology, law] means the back catalogue is of considerable importance.”

Archiving of digital and electronic materials is a current area of research—and debate—among librarians, archivists, and information and computer scientists. In fact, as Arms (1999b) points out, there is not even agreement that such material should be archived since “publishers and librarians often equate primary information with conventional peer-reviewed journals,” but “practicing scientists recognize that that is far from accurate.” The Task Force on Digital Archiving and the Council on Library and Information Resources provide good introductions to the issues in this area.⁴⁵

Electronic/digital archiving entails a series of organizational and technical challenges, beginning with selection of materials and definition of requirements. Is the entirety of an issue of an online magazine to be preserved? Is it sufficient to create a database of individual stories that can be individually retrieved but never reconstituted as a specific issue of a newspaper as it may have existed on the day that a reader first read the news?⁴⁶ Additional questions entail responsibility (publisher versus library) for long-term preservation, copyright, and maintenance; these issues arise from the risk that, over time, the technologies required to view a given file may

become obsolete, effectively eliminating the record.⁴⁷ Arms (1999b) suggests a two-step process in which a period of active management is followed by preservation of the original by the publisher or designated entity (e.g., HighWire, JSTOR, or some similar archiving service).

MEASUREMENTS OF IMPLICATIONS AND CHANGES IN RESEARCHER BEHAVIOR

“Impacts” of individual articles or journals are typically measured through bibliometric techniques, which are well understood among information scientists. There are efforts to extend this methodology to hyperlinks (i.e., the links in an online document that lead either to another site or to another point in the same document) and also to extend the notion of “citations” to include broader acknowledgments reflective of other ways that impacts and conceptual relationships might be expressed. This appears to be particularly interesting in the Web environment, which provides multiple forms of offering information (Web pages, white papers, electronic versions of peer-reviewed journals articles, listserv discussions, and so on) as well as the means for establishing explicit connections through hyperlinks. Formal assessments of the impact of electronic publications and expanded methodologies for conducting these studies are in their infancy and appear to be very promising.

Broadly speaking, two kinds of studies look at measurements of implications and how these are reflected in the scientific scholarly literature: bibliometric studies and behavioral studies. Bibliometric studies examine patterns of citations and other references within the literature as a measure of the relative importance of a given article. Behavioral studies use observation, focus groups, questionnaires, and interviews to examine people’s actions and attitudes. We here distinguish between behavioral studies and user studies, the latter of which are focused on designing and evaluating systems rather than on assessing the significance of research, although clearly

⁴⁵Task Force on Digital Archiving, *Preserving Digital Information: Report of the Task Force on Archiving Digital Information*, study commissioned by the Commission on Preservation and Access and the Research Libraries Group, Inc., <http://www.rlg.org/ArchTF/tfadi.index.htm> (1996); Council on Library and Information Resources, *Authenticity in a Digital Environment* (Washington, DC, 2000).

⁴⁶It is essentially impossible to reconstruct the Web versions of *The New York Times* and *The Washington Post* on a daily basis, although individual stories are accessible from the papers’ online archives.

⁴⁷This is a well-known problem for public records. See Henry Gladney, “Archiving the Digital Public Record: An Internet Snail’s Pace,” *iMP: The magazine on information impacts* October 2000, http://www.cisp.org/imp/october_2000/10_00gladney-insight.htm; and Raymond Lorie, *The Long Term Preservation of Digital Information* (IBM Almaden Research Center: 2000), <http://www.almaden.ibm.com/u/gladney/Lorie.pdf>. In this context, too, software obsolescence is another key concern for researchers and managers of digital archives.

there is overlap. Results are described in chronological order based on the date or period in which the data were collected or, in the cases of citation analysis, the period covered by the research. In both broad classes of research, definitions of the notion of electronic publishing—and consequently the scope of coverage—changed over time, and results are thus not strictly comparable. In general, electronic journals are witnessing greater acceptance among researchers although there remain concerns over prestige, archiving, and the stability of the record.

BIBLIOMETRICS: CITATION ANALYSES, ACKNOWLEDGMENTS, LINKS, AND INVOCATIONS

According to Borgman (1990a, p. 11), “Bibliometrics encompasses a number of empirical indicators that can be found in the formal record of scholarly communication, including authors, citations, and textual content.” As she points out, computing technologies that enable the creation of text and databases in machine-readable form, together with statistical application software, have heightened the scale and scope of bibliometric analysis. The essays in Borgman (1990b) provide an introduction to the analytical issues and methods entailed in bibliometric analysis in the field of scholarly communication.

Several studies on the implications of electronic publishing have focused on some form of citation analysis. Table 2 summarizes key characteristics of these studies;

note that the definitions and scope vary from study to study. Some researchers have made efforts to extend their analysis to patterns in hyperlinking and to new ways to understanding impacts that are appropriate to the Web environment. In this regard, Treloar (1996) observes that hyperlinking “provides one-way (source to specified target) links only.” That is, a hyperlink reflects relationships from the article to the more general literature; it does not reflect references from the literature to the story, a pattern that analysis of citation databases can support (Treloar 1996). On the other hand, relying on citation index databases effectively limits the population on which the assessment is based to the journals covered by the indices and excludes, for example, Web pages, preprints, working papers, and technical reports (Harter 1996).

Harter (1996) conducted a citation analysis of the 39 peer-reviewed scholarly electronic journals published no later than 1993—that is, before the expansion in e-publishing associated with the launch of the World Wide Web. The unit of analysis was the journal (rather than an author or a given article) since “the total number of citations tends to measure the cumulative effect of a journal...It is a long-term measure, clearly biased in favor of older journals that have been publishing for many years” (Harter 1996). Harter found that, in general, e-journals were not frequently cited; he identified the top eight cited journals as being in the fields of (in descending order of citation frequency) mathematics, medicine, library and information science, computer science, psychology, effects of technology on society, communication, and

Table 2. Summary characteristics of citation analysis studies

Study	Year(s) covered by data	Unit of analysis	Sample size	Field	Scope
Harter 1996, 1998	1993 and earlier	Journal	39	All fields	Peer-reviewed e-journals, including journals in parallel forms ¹
Youngen 1997	1988–96	Journal	Not given	High energy physics, astrophysics	Citations to preprints in print and digital form from the journal literature
Zhang 1998	1994–96	Article	14 journals	Library/information science	Citations to electronic sources including but not limited to e-journals from print and e-journals
Hurd, Blečić, and Vishwanatham 1999	c. 1995	Article	44	Molecular biology	Publishing choices and behaviors of a specified faculty including their use of electronic resources and journals
Cronin et al. 1998	c. 1997	Individual	5	Not specified	Development of a typology of kinds of signaling and linking behavior in the context of the Web
Kaplan and Nelson 2000	1997	Technical report	50	Aerospace engineering	Technical reports contained in the Langley Digital Library, use by the community and citation in the literature

¹Harter (1998, p. 507) defines e-journal as follows: “using computers and communication networks to create alternative electronic forms of the conventional paper journal.”

modern culture. He then compared these findings with a contextual analysis of journals in these fields, noting, however, that citation patterns within and among disciplines vary, so that it is “unfair and misleading” to make statements about author productivity and journals in different fields. That said, he found that, overall, “most e-journals are having little impact on formal scientific and scholarly communication.”

Harter then considered two other metrics: the impact factor (which looks only at the 2 previous years of publication, thus controlling for the effect of time) and the immediacy index, which considers the extent to which articles make a “quick impact” on readers. Two of the eight most highly cited e-journals existed in parallel forms (print and digital), which conflated the measurement. This resulted in identifying seven “very influential” articles, of which four were in medicine (including the top three) and three in electronic publishing and computing. Harter notes in his conclusion that one limitation arose simply because there were fewer articles in e-journals and that “authors will need to view e-journals as legitimate publication vehicles before e-journals can assume a significant role in the scholarly communication process.” Harter (1998) reinforces this conclusion in an article in which he expands on these findings, including responding to comments received after the release of the previously described study in the online *Public-Access Computer Systems Review*.

The thrust of the research conducted by Cronin et al. (1998) is that a hyperlink is broader than a formal citation and can include reference to underlying evidence or, more casually, to listservs and discussion fora. These latter capture the notion of “acknowledgment,” which demonstrates “peer interactive communication.” Thus, Cronin et al. posit the term “invocation” (pp. 1319, 1320). The goal of their study was to develop a typology for this new citation form, which they did by conducting a survey based on a sample of five individuals in library and information science, selected from a larger population of highly cited library and information science faculty. The names were submitted to five then-prominent Web search engines over a period of 2 months in 1997, and it was found that all five individuals had personal home pages, making this the most frequently identified typology for the group. The next most frequently identified typology was “conference proceedings,” suggesting to the authors that the Web could help elevate conference proceedings to “a higher level of visibility” (p. 1324). In general, the authors conclude that, in contrast to formal

citation analysis, which “can tell us a lot about the formal bases of intellectual influence,” structured analysis of the Web offers information “about the many other modalities of influence which comprise the total impact of an individual’s ideas, thinking, and general professional presence” (p. 1326).

Four studies examined information use through citation analyses in molecular biology, physics, and library and information science. Youngen (1997) and Kaplan and Nelson (2000) focused on the literature in digital form; Hurd, Blecic, and Vishwanatham (1999) and Zhang (1998) considered e-publications in a larger media-independent context.

Youngen (1997) conducted a citation analysis of the physics preprint literature based on the Internet-accessible collections at Los Alamos National Laboratory (LANL) and the Stanford Linear Accelerator Laboratory (SLAC). His paper, which was given at a professional conference in June 1997, offers a concise introduction to the history and structure of these important collections as well as to some of the methodological issues that arise in conducting such a study. Youngen distinguishes between *preprints*, which are “manuscripts that are intended for publication but are being circulated among peers for comment prior to being submitted for publication” and “are likely to be the earliest version of a study,” and *e-prints*, which are preprints in electronic form that have been assigned a standardized identification number upon their submission to LANL or SLAC. Preprints may exist in either print or digital form; e-prints are only in digital form. Youngen finds that citations to preprints declined consistently from 1988 to 1996 but that citations to e-prints increased rapidly since their introduction in 1992. This finding suggests that e-prints are becoming more accepted within the physics and astronomy communities and by publishers and editors and that scientists working in subjects in which preprints are common are “making the transition to electronic publications.”

Examination of the relationship between preprints/e-prints and journals indicates that the e-print phenomenon is “most strongly present in high energy physics (particularly particle physics) and astrophysics.” Youngen notes that there may be conflating and biasing effects resulting from his methodology and editorial policies that encourage or discourage use of the e-print identifier (e.g., by preferring language such as “submitted to” over the term “e-print”).

Zhang (1998) conducted a statistical analysis of articles in 14 peer-reviewed/refereed journals in library and information science for the period 1994–96 and asked three questions: (1) What is the overall impact of Internet-based e-sources on formal scholarly communication?, (2) Is there evidence of increasing impact during this period? and (3) Are e-journals more likely to cite electronic sources than print journals? He concluded that the impact of e-sources on formal scholarly communication in library and information science is much smaller than that of print, as measured by e-sources cited in all journal articles published in the period under study, although there was an observable increase that was not statistically significant (based on a one-way analysis of variance) at the 0.05 level. Zhang also notes that the pattern of citation indicated a turning point in 1994 (p. 249). There is evidence that e-journal articles are more likely to cite e-sources than print journal articles, but there is no significant difference in the number of references per article by journal format once an author cites e-sources. Zhang concludes “that citing e-sources may depend on authors rather than the journal format in which the authors choose to publish their work” (p. 249). He also notes that variation in citation conventions appears to be one barrier to universal citation of e-sources; a second is the dynamic nature of the Internet, which may lead to concerns over the permanence and continued accessibility of the cited source, editorial policies on archiving notwithstanding (p. 249).

Hurd, Blečić, and Vishwanatham (1999) examined information use by molecular biologists. Their principal interest involved issues in library collection management and acquisitions, but their results are pertinent to how molecular biologists view and use electronic journals. The study comprised a citation analysis based on the authoring patterns of 24 faculty members in five different departments at the University of Illinois at Chicago in 1995. Collectively, these academics published 60 articles; 44 of these articles, published in 27 journals in biology, medicine, and science, were included in the final sample (no more than three articles per faculty member were included, and 20 of the 24 academics had published at least one article within the last 3 years).

None of the articles in the sample of faculty publications cited any electronic source, whether an electronic journal or database; four articles indicated that genetic sequences determined in the research reported were available for reader access in a computer repository. Three authors identified deposition at GenBank or the European Molecular Biology Laboratory database, and

one author stated that the data were available via a personal FTP (file transfer protocol) site. The authors point out that although various kinds of indexing, searching, and access aids had been available online or on CD-ROM (e.g., Medline, BIOSIS) for 20 years, electronic journals were relatively rare in biology at the time of the study; the genome databases were also relatively new (p. 40).

Harter’s 1996 study has some bearing in this context, as four of the seven most highly cited e-journal articles he identified were in medicine. Harter’s study is roughly contemporaneous with the period covered by Hurd, Blečić, and Vishwanatham, during which they found that molecular biologists did not publish in or cite electronic sources and were, as a group, only slightly involved with the genome databases. The studies differ methodologically; nevertheless, the diametrical contrast in their findings suggests that different methods may yield significantly different results, and/or that there are substantial differences in the publication (and hence the communication) patterns within the life sciences as well as between the physical and life sciences and between the natural and social sciences.

Finally, Kaplan and Nelson (2000) analyzed technical reports in digital form covering aerospace scientific and technical information contained in the Langley Technical Report Server for the year 1997. This material was compared with print materials managed by the Center for Aerospace Information (CASI). The investigators identified the 50 most frequently accessed reports in electronic form; they then conducted a citation analysis of these reports, based on the journal articles indexed by the Science Citation Index and the Social Science Citation Index. This meant, the investigators note, that the examination considered the “impact” of technical reports, which are made available to the professional community but are not formally published in the peer-reviewed journals. Based on this analysis, they found almost no impact, nor was it clear whether the citations were to the electronic version available via Langley or to the print version available through CASI (p. 329). Noting that “uncitedness does not equal useless” (p. 331), Kaplan and Nelson propose a new metric, retrieval analysis (i.e., how people use the information they find). Comparison of the electronic collection with the print collection showed that the most frequently requested item from the print collection was requested 45 times, but that the top item in the digital collection was accessed 672 times, leading the investigators to suggest that the electronic format “contributed to its higher usage numbers. Users can browse, search, locate, download and use items

immediately” (p. 332). This conclusion was reinforced by comparisons at the collection level. The digital library had holdings of approximately 1,400 items and distributed approximately 71,000 copies in 1997. In the same period, CASI contained holdings of 3.5 million items and distributed approximately 24,000 copies (p. 334).

STUDIES OF BEHAVIORS AND ATTITUDES

We identified 12 surveys covering the 1991–99 period of how authors and readers use electronic journals; the characteristics of these studies are summarized in table 3. Despite the diversity in methodologies and approaches, the surveys collectively suggest greater use and acceptance of electronic journals over the period, particularly in mathematics and natural science, where there are more e-journals and a somewhat longer tradition of use.

The surveys rely almost entirely on questionnaires, sometimes amplified by interviews and, in one case, by bibliometric and computer log analysis. How the samples were drawn (whether by devising a formal sampling strategy or identifying a target population of, for example, faculty in different disciplines) varies among the surveys, as do the range of questions posed, definitions (including definitions of an electronic publication) used, variables examined, and study size. Our research also identified two studies based on the SuperJournal project (Pullinger 1999 and Yu and Apps 2000); these examined e-journals made available by European publishers in the context of information-seeking behaviors.⁴⁸

Hurd and Weller (1997) offer the results of two surveys (in 1991 and 1995) on the adoption of information technology by faculty at the University of Illinois at Chicago.⁴⁹ Although all faculty in the basic sciences, health sciences, and engineering were covered in the surveys, the paper only reports on data from chemists and chemical engineers, whose response rates were higher than for the entire population of faculty in both surveys. Between 1991 and 1995, access to the university’s network increased: 30.2 percent of the chemist/chemical engineer faculty reported no access in 1991; 4 years later, 100 percent had active accounts. Use of various electronic indexes (e.g., the Chemical Abstracts Service, Indicus Medicus, and MEDLINE) generally increased over the period, with variation based on faculty subspecialty. Use of MEDLINE and ISI Current Contents, both

of which were then available via the university network to faculty member offices, showed high growth rates.

The authors did not find a clear relationship between a given faculty member’s subspecialty (which might indicate his or her computer literacy) and his or her use of these resources, leading them to suggest that convenience was the driver of use. When respondents to the 1995 survey were asked about their Internet/Web usage, about half said that they used different functions; there was again some variation observed among faculty subspecialties.

Schauder (1994) surveyed 743 senior academics in Australia, the United States, and the United Kingdom in July 1992, taking the names of the individuals surveyed from the 1991 *World of Learning* directory. Respondents agreed that “prestige” and “readership” were the determining factors in their selection of journals in which to publish; and, although respondents believed that time to publication should be quick (3 to 6 months), they were willing to accept longer time frames. They were similarly positive about peer review and “not very concerned about its perceived weaknesses.” There was general agreement that serials were too expensive, particularly among respondents from the physical sciences and engineering, followed—in descending order of number of respondents agreeing—by those in the biological sciences and medicine, social sciences, and law. Only 39 percent currently used a national or international network, but those who did were disproportionately concentrated in the biological sciences and medicine; followed by the physical sciences and engineering; and then by the social sciences, law, and business. The most popular Internet application reported by respondents was e-mail; use of listservs or newsgroups for either obtaining or publishing articles was “very much a minority activity” (p. 89). Respondent use of information technologies was widespread but was distributed over four activities: locating information, filing or organizing information needed for research and writing, collaboration with others either in writing or informal review, and preparing drafts.

When asked whether e-publication would obtain the same recognition in promotion and tenure decisions, about equal proportions responded “yes” and “don’t know” (35 and 33 percent, respectively). Only 12 percent, however, responded “no.” Asked whether users should be allowed to print or copy to disk articles that they (the respondents) had written, a large majority (78 percent)

⁴⁸These studies have not been included in table 3 since they employ multiple methodologies.

⁴⁹These surveys did not focus on electronic journals per se.

Table 3. Summary characteristics of surveys

Study	Year(s) in which data were collected	Sample size	Response rate	Field or target group	Scope of Study
Hurd and Weller 1997	1991 1995	53 54	66.3%, 63.5%	Chemistry and chemical engineering	Print and electronic versions of various finding aids (e.g., the Chemical Abstracts Service) and the Internet, including listservs, e-mail, the Web, FTP, gopher, and Telnet
Schauder 1994	1992	743	78%	Senior faculty in Australia, the U.S., and the U.K.	"Publishing of professional articles in electronic form" (p. 73.); includes availability of the full text of articles via electronic databases; "electronic publishing" is defined as "(a) dissemination and archiving via computer storage media, and (b) access through computers in stand-alone or network mode" (p. 94); these conditions were met by "a large and growing number of journals which are parallel published through online hosts and on CD-ROM"
Berge and Collins 1996	1994	256	16%	Computer science	Readers of <i>Interpersonal Computing and Technology: An Electronic Journal for the 21st Century (IPCT Journal)</i>
Budd and Connaway 1997	1995	651	48%	Chemistry, physics, sociology, psychology, English, history	Networked information, including e-journals (not defined) and Internet-accessible bibliographic databases and library catalogs
Kaminer and Braunstein 1998	1995	122	52%	Departments of Agriculture and Resource Economics, Environmental Science, Policy and Management, Nutritional Science, and Plant Biology	Internet use, including e-mail, Web, listservs, and e-journals (not defined), as a means of modeling scholarly productivity
Gomes and Meadows 1998	1996	213	14%	Faculty in the basic sciences (physics, chemistry, and biology) in the U.K.	E-journals, whether pure or parallel with print not stated as an element of the design, but the authors observe that the existence of a print version may affect perceptions of prestige (p. 180)
Tomney and Burton 1998	1996/97	75	28%	Faculty across disciplines, including science, engineering, arts, business, and education (U.K. universities)	E-journals, defined as one that publishes original scholarly writing and is available, not necessarily exclusively, in electronic format (p. 420)
Hamerslag 1998	1996	169	33%	Medicine (Israel)	E-journals, whether pure or parallel with print not specified
Speier et al. 1999	1997/98	300	22%	Business school faculty	E-journals, whether pure or parallel with print not specified
McKnight and Price 1999	1997/98	537	52%	Journal authors in science, arts, and humanities	E-journals, whether pure or parallel with print not specified
Brown 1999	1998	49	61%	Astronomy, chemistry, mathematics, physics	Departmental library collections including books, monographs, journals, preprints, conference proceedings in electronic and print forms, on site and via networks; definition of e-journal not provided
Rusch-Feja and Siebeky 1999a, b	1999	1,042	11%	Biomedical, chemical-physical-technical disciplines, humanities	E-journals made available by major European publishers

answered “yes,” although the responses varied somewhat when users were defined as either university-based, corporate, or commercial users. Still, respondents generally favored conditions easing access by end users. Finally, when asked whether their university was currently active in e-publishing, 62 percent responded “yes” or “to some extent.” When asked if a university should be engaged in e-publishing, 30 percent responded “don’t know,” and 53 percent answered either “yes” or “to some extent.” Interest in electronic publishing was higher among younger academics and those who were already network users.

Berge and Collins (1996) examined issues related to e-journal author and reader concerns, as well as to the acceptance of e-journals for promotion and tenure, through a 1994 online survey of the readers of *Interpersonal Computing and Technology: An Electronic Journal for the 21st Century*. The study yielded demographic information and generally positive responses to the quality of the journal; it also found a high level of acceptance by the readership, although acceptance for purposes of author promotion and tenure was considered “largely problematical.” The investigators were surprised by readers’ relative lack of concern for permanence; they caution that the survey method and low response rates may have self-selection effects on the findings.

Budd and Connaway (1997) looked at current practices and attitudes of university faculty toward networked information as it affects academic work. They conducted a survey of university faculty in six departments—chemistry, physics, sociology, psychology, English, and history—at eight geographically dispersed universities in the United States in the fall of 1995. The questionnaire covered several dimensions of communication and information seeking. Overall, only 14 percent of the respondents used electronic journals, although the usage rate varied substantially by discipline. For example, while 6 percent of the chemists surveyed subscribed to at least one electronic journal, 32 percent of the physicists did so; respondents in English, history, psychology, and sociology departments were close to the aggregate response rate. Reported usage was higher when all electronic sources were considered: 64 percent of the sociologists, 56 percent of the physicists, 52 percent of the chemists, and 32 percent of the psychologists reported using some form of electronic source material. Only 32 percent of all respondents believed that university tenure and promotion committees were open to accepting e-publication, although 40 percent believed that there had been a change in attitudes toward electronic publication. Respondents

also reported using the Internet to find out about sources of funding; this was particularly true for younger faculty members.

Only 3 percent of the respondents reported submitting articles to electronic journals. The authors note, however, that, at the time of the study, there was only an extremely small number of e-journals in which to publish (p. 675). Survey respondents commented that peer review and a publication’s prestige were critical factors in their decision to publish in a given e-journal; they also noted that these journals’ (lack of) permanence was an issue of concern. Commenting on research previously conducted by Schauder (1994), Budd and Connaway note that there is a tension between the perception of prestige associated with journals with a high price and readership, which is affected by pricing but which is also a factor in selecting journals in which to publish. The authors conclude that the present system of promotion and tenure favors traditional means, thus hindering adoption of the innovation—i.e., e-journals.

Kaminer and Braunstein (1998) consider the use of the Internet in their model of scholarly productivity, using data from 1995. Based on a survey of 122 faculty members at the University of California–Berkeley’s College of Natural Resources, a bibliometric study based on bibliographies maintained by the Academic Personnel Office, and analysis of computer logs, the investigators conclude that the effect of the Internet, broadly defined, can be quantified and compared with the use of more traditional forms of information. They find that Internet use has a positive effect on scholarly productivity as measured by publication rate, weighted by such considerations as author age, years since acquiring doctoral degree, etc. Given the period in which the study was conducted (1995), it is telling that the dominant Internet use was for e-mail (94 percent), followed by Telnet (62 percent)⁵⁰ and the Web (44 percent). Electronic journals were used by only 9 percent of the respondents; this term was not defined in the survey.

Gomes and Meadows (1998) conducted a study of staff in British university science departments in 1996 to look at the general question of how any given scholarly community would adopt electronic journals. Of the 1,480 questionnaires sent out, 213 were returned, mostly

⁵⁰Telnet is “A terminal emulation protocol that lets a user log in remotely to other computers on the Internet...” (Computer User High-Tech Dictionary, <http://www.computeruser.com/resources/dictionary/index.html>).

from physics and biology staff, and fewer from chemistry. The sample also included faculty in library and information science; the authors did not tabulate the responses by discipline. Nearly all respondents (95 percent) used e-mail at least once a week. They cited the most important characteristics of journals—either print or electronic—as being the quality of the articles published and the prestige of the journal. Respondents who published frequently were more concerned with the prestige of the outlet than with cost; those who published infrequently were more concerned with cost. Authors were more likely to publish in electronic journals if these were accepted as appropriate outlets by their respective research communities, by funding agencies, and by promotion and tenure review committees. Librarians more or less echoed the concerns of their scientist colleagues, but expressed greater reservations about preservation (i.e., archiving). The survey found that university administrators appear to have no explicit policies concerning the acceptability of electronic publication for promotion and tenure, and only 14 percent of those surveyed believed that articles in e-journals were unlikely to be accepted for such review by the end of the 20th century. The major sticking point regarding the acceptance of electronic journals, the authors conclude, was the perceived lack of prestige.

Tomney and Burton (1998) also assessed attitudes toward electronic journals based on a stratified sample of individuals from 10 departments in an unnamed British university. The authors do not give the timing of the study, but internal evidence suggests that it took place in 1996/97. Slightly less than one-third (28 percent) of the respondents reported using electronic journals; usage rates were highest among the business, science, and engineering faculties. In general, both readers and nonreaders of e-journals used other electronic resources, suggesting that “familiarity does not necessarily encourage or promote” the use of e-journals. Nonusers of e-journals across all departments, except physics/applied physics, cited “lack of awareness” as their reason for nonuse; only 13 percent gave the reason “not rated as highly as printed journals” to explain their nonuse. The most frequently cited advantage of e-journals was accessibility. Hypertext links were seen as a clear advantage to users, but the ability to attach comments was not; nonusers, however, saw both features as advantages. The three greatest disadvantages to both users and nonusers were copyright, the potential for text alteration, and the impression that an electronic publication is not a “real” publication. Based on the survey results, the authors suggest that the barriers to acceptance of e-journals are time

and comfort with the technology; “for a large number of academics (of all ages and grades, although predominantly among the more senior positions), they simply feel that they do not have the time or experience to search for these electronic journals on the Internet” (p. 427).

Hamerslag (1998) describes the results of a survey of Israeli medical researchers—faculty at Tel Aviv University, Ben Gurion University, and the Israel Institute of Technology—that took place in 1996. While respondents were aware of some of the advantages of electronic journals, they read them infrequently and used them as a publishing tool to a very small extent. Most respondents doubted their value for academic promotion. Hamerslag ascribes the low interest in reading to the small number of e-journals, lack of recognition of their scientific value, conservatism, inefficient search practices, and the inconvenience associated with reading on screen.

Speier et al. (1999) looked at the perceptions of electronic journals among business school faculty. Overall awareness of e-journals was extremely low among the population surveyed, and less than 10 percent intended to submit or had submitted articles to electronic journals. More prolific faculty who had already obtained tenure and who served on promotion and tenure committees were more likely to submit articles to e-journals. Faculty in finance and in management of information systems were more likely to read articles in electronic journals than were their colleagues in marketing, operations, management, etc. Overall, however, e-journals were not perceived by the respondents as being of as high quality as their print counterparts.

McKnight and Price (1999) examined various aspects of authors’ experience, attitudes, and perceptions of publishing in paper and electronic journals, based on a 1997–98 survey of 1,040 British authors in science and the arts and humanities. The response rate was highest among scientists, which may be associated with the presence of online science journals. Nearly all survey respondents had published within the last 2 years, but only 16 percent said that they had published in an electronic journal. Interestingly, fewer science authors had published in e-journals than had their peers in the arts and humanities. The authors note, though, that during this period, the monthly submission rates to the LANL preprint archive increased rapidly. Respondents were generally positive about peer review (science, 94 percent; arts, 96 percent; humanities, 93 percent), although there was some confusion over the nature and extent of peer review among

electronic journals. Awareness of electronic journals was generally high, and over three-quarters of the respondents believed that e-journals were “here to stay.” Use of multimedia was attractive to about a third of the respondents: although information technologies had achieved almost complete penetration of the academic workplace, the skills required to generate multimedia submissions appear to be a deterrent to their use. The authors speculate that the more general deterrent to publication in e-journals is their perceived transience and low prestige.

Brown (1999) considered the information-seeking behavior of astronomers, chemists, mathematicians, and physicists at the University of Oklahoma. In response to a survey question on resources used for research, chemists, astronomers, and physicists indicated that journals were their primary source; mathematicians also indicated a reliance on preprints, conference attendance, and personal communication. All respondents relied on textbooks for teaching. The faculty surveyed used various mechanisms to ensure current awareness in their field, including scanning tables of contents and electronic “current awareness” services. The mathematicians showed a reliance on MathSciNet (a comprehensive database of mathematical literature maintained by the American Mathematical Society), and the physicists and astronomers listed the LANL and SLAC collections as important databases for current awareness and research activities. The majority of respondents preferred to access journals in print rather than in electronic form; those who preferred access in both media indicated a desire to print out articles, even if they were initially delivered in electronic form.

Rusch-Feja and Siebeky (1999a, 1999b) report the results of a survey of the Max Planck Society which was conducted in April and May 1999. The society encompasses 84 research institutes, with several additional working groups and research centers covering the full range of disciplines from physics and chemistry to the humanities. The study distributed a questionnaire, and compared the results with information obtained from major commercial publishers engaged in e-publishing (Elsevier, Springer, and Academic Press). In all, 1,042 valid sets of answers were received, of which 50 percent came from the society’s biomedical section, 38 percent came from chemical-physical-technical disciplines, and 12 percent came from the humanities.

Most of the survey respondents said that they used e-journals every 2 weeks to once a month. Elsevier journals were used most frequently, but, as the authors observe, Elsevier also offers the greatest number of journals by a factor of more than two. The biomedical researchers were the most frequent users; the next most frequent set of users was chemical-physical-technical researchers, who reported using the relevant journals once a month. This level of frequency for science users reflects the availability of journals online as well as the importance placed on currency of information. Not surprisingly, given the results of other surveys and user studies, the advantages cited for e-journals included direct access from the desktop, prompt availability, ability to download and/or print the information, and currency. Disadvantages cited were lack of long-term access and the assurance of archiving and the incompleteness of the backfiles.

Yu and Apps (2000) review the issues related to log file analysis and describe the methodology used to overcome these in support of the SuperJournal project (see sidebar on the following page). They define log files as “text files generated by certain software packages to record events and the time they occur in a computer system” (p. 311). Analyses based on these data have been a feature of information retrieval studies for several decades, but gained widespread use with the advent of the Web in the 1990s. Log files pose challenges for statistical analysis because of their very large size as well as because of ambiguities introduced by the nature of the underlying communications architecture. In addition, log files for e-journals pose particular problems related to user privacy and interpretation of behavior—how to discriminate, for example, between browsing and close reading.

Pullinger (1999) reports the results of part of the SuperJournal project, a 1996 survey of 70 scientific users in four universities, 94 percent of whom used the Internet at least weekly and 80 percent of whom used bibliographic databases at least weekly. The survey respondents used a variety of online journals and services, of which the four most important (in descending order of priority) were *Nature*, *Journal of Biological Chemistry*, *Cell Science*, and Bath Information and Data Services (BIDS—<http://www.bids.ac.uk/>).

He then describes some of the preliminary results of the larger project, based on log files for 1,817 registered users as of August 1998. Among regular users, scientists

The SuperJournal Project

The U.K.-based super journal project studied the factors that will make electronic journals successful. It involved collaboration among a group of 17 learned society, university, and commercial publishers; the University of Manchester; and the University of Loughboro. The project ran from 1996 to 1998 and was eventually deployed in 13 universities. The summary report for the SuperJournal project as well as numerous technical reports and supporting papers are available at <http://www.superjournal.ac.uk/sj/index.htm>. The project employed a variety of methodologies, including questionnaires, focus groups, and Web log analysis. Users were divided into broad categories, which included “scientists” and “social scientists.” Unfortunately, the sizes of the groups are not always stated in the available documents, and, although there were followup studies, it is unclear whether these were in fact formal longitudinal cohort studies (i.e., that they looked at the same people at successive intervals). One of the strengths of the project was its attention to baseline and followup studies in order to capture changes: the investigators clearly and intentionally set out to distinguish between what users said they wanted both at the outset and farther into the project, what their behavior revealed, and how this changed or appeared to be inconsistent. The project’s key findings include the following.

- The most important requirements for electronic journal services are a critical mass of journals, access (particularly at the desktop), and timeliness. Science users rated this last requirement as the most important; social science users rated the range of journals as being most important. The number of relevant journals is the most important factor in determining whether a user will use an e-journal service. Electronic journals were seen as a way of increasing access to the professional literature. In the 1996 baseline

studies, half of the science readers and 70 percent of the social science readers said that their access to journals significantly limited the breadth and depth of their reading; they cited a lack of journals in the library as the most frequent reason for this limitation. In a similar survey at the end of the project in 1998, over 90 percent of the social science readers said they used journals in SuperJournal that they had not read before, and 70 percent of them used journals that were not available in the library. Users want seamless access to important collections of relevant material, whether bibliographic databases or collections of related journals.

- Core functionality is another top requirement for an electronic journal service: users must have the ability to browse, search, and print. The ability to search the service appears to be more important to social scientists, who are not as well served as scientists with bibliographic databases in their disciplines. Printing is valued because users generally do not like to read on screen, preferring to take material with them to more comfortable quarters where the text can be annotated. Users expressed concern about changes that might damage the quality of articles, in terms of both content (peer review) or format, or endanger the permanent archival function of journals.
- Users also cite the need for an electronic journal service to have a backfile, ideally of 5 to 10 years; this requirement is probably most important to those in the social sciences. As important as access to the backfile is the knowledge that the journal content (both current and back issues) will remain available into the future and will not disappear (e.g., when a subscription ceases). Depth, permanence of content, and breadth of the collections enhance the overall critical mass, which encourages users to make more extensive use of the service.

accessed it on average once a month and social scientists less frequently, 2 out of 3 months.⁵¹ More frequent use by scientists, Pullinger suggests, may be due to the relevance of some journals to researchers’ core interests vis-à-vis SuperJournal’s offerings, which may be of interest but peripheral relevance. SuperJournal was generally used outside of conventional office and library hours. Social scientists seemed to access far more journals than did scientists, suggesting that the latter are

“much more specific” in their selection of journals (p. 170). Similarly, searching was more highly engaged in by social scientists, but regular users “browse” more (that is, skim or scan a document) and search less (that is, submit a formal query through, for example, a search engine or service). Pullinger suggests that the focused use of certain journals by scientists may indicate that their searching is done elsewhere, and that they are using the online journals as a “current awareness checking procedure or for access to particular papers when they had already been identified as relevant” (p. 171).

⁵¹The SuperJournal project distinguishes between “scientists” and “social scientists.”

IMPLICATIONS FOR UNDERSERVED POPULATIONS

Electronic journal publishing can contribute to greater access to scientific research and communication for those groups in the United States and abroad that are currently less able to acquire printed journals or travel to professional meetings because of limited financial resources and other constraints. The following material examines what the literature says about the ways and circumstances in which e-journal publishing can offset some of these disadvantages and enhance access of underserved populations to scientific findings and conversations. But despite widespread interest in the potential of the technologies to offset disadvantages, there is relatively little research that focuses specifically upon the scholarly journal and the global scientific community.

Cohen (1996) describes a debate among scholars of computer-mediated communication concerning the potential democratizing effect of these systems. His own study, however, yielded no data to support such an effect, at least among research universities and master's level institutions. More recently, studies of the so-called digital divide⁵² (NTIA 1999) have raised concerns about, and heightened visibility of, the issue of universal access to computers and information technologies in the United States. Some experts and policymakers fear that those without access to computer networks and without relevant computer literacy will find themselves increasingly at a disadvantage. In response, one feature of the Telecommunications Reform Act of 1996 was assistance to public schools to enable them to obtain Internet access.

The American Association for the Advancement of Science (1998) has called attention to the potential opportunity that electronic publishing offers for developing nations to promote the advancement of their scientific information (see also ICSU Press/UNESCO 1996). Schoonbaert (1998, p. 101) echoes this idea and addresses the implications of electronic publishing in the context of access by developing nations to medical and public health information. He comments that such access "is an enormous step forward...as seen from the spectacular improvement of disease outbreak information" and speculates that although affordable pricing is a concern, timely

access to at least portions of the literature (e.g., tables of contents, abstracts) "certainly helps" (p. 102). Information exchange travels both ways: "the electronic media also enables Third World research to profile itself on a global level" (p. 102).

E-publishing and e-journals are making inroads in some countries. Although Bavakutty and Radhamani (1991) found that the dominant mode of sharing and disseminating scientific research in India based on library practices was print, their study was conducted over a decade ago, and access patterns have almost assuredly changed in the interim. More recently, Pakenham-Walsh (2000) notes that the African Journals Online project, which was set up in 1997 with funding from UNESCO, has apparently succeeded in raising awareness of African journals in science. Adam (1999) concludes, however, that substantial basic infrastructure investment is required to ensure affordable access as a precondition to promoting scientific communication and publishing in Africa.

The Scientific Electronic Library Online is an electronic virtual library covering a selected collection of Brazilian scientific journals (<http://www.scielo.br/>). The library is an integral part of a project being developed by the State of São Paulo Research Foundation, in partnership with the Latin American and Caribbean Center on Health Sciences Information. According to its website, the project "envisages the development of a common methodology for the preparation, storage, dissemination and evaluation of scientific literature in electronic format." As of this writing, the site lists 57 journals in English and Portuguese.

Some efforts are being made to address the pricing issues that affect access. Rous (1999) reports that the Association for Computing Machinery (ACM) discounts consortia prices based on the size of the consortium and the U.N.'s three-tiered classification of countries by wealth. As a result, ACM has seen its global reach expand. Peters (2000) describes the considerations behind the pricing strategy for *Sociological Research Online*, which was intended to ease access for independent scholars worldwide and for institutions in developing countries. Specifically, free subscriptions would be given to countries whose GDP falls below a given threshold. Peters sets forth a model for access and charging that accommodates independent scholars, including those in emerging nations.

⁵²This term is typically used in describing the phenomenon of those people who do and do not have access to computers and the Internet.

The question of moral rights is emerging at the global level as a “hot button” issue for emerging nations that appear to be suspicious of the equity of the intellectual property rights’ regimes of the developed nations. To date, this has not emerged as a discussion specific to scientific and scholarly journal publishing.

INFORMATION SECURITY AND USER PRIVACY

Information security involves policies, procedures, practices, and systems intended to protect and authenticate information and data in all its forms, including storage, transmission, and use. This topic includes information on individuals’ use of system and personal data that may be stored and collected. For our purposes, information security systems include means of detection as well as methods of maintenance and control. This is another area in which there is interest but relatively little research.

Krieb (1999) provides a description of the experience of Saint Louis University’s Health Sciences Center Library in providing access to dispersed users, including attention to information security at the system level. Arms (2000) links issues of access management (i.e., who has access to materials and for what purpose) to security techniques in networked computing, noting that “when publishers expect revenue from their products, they permit access only to users who have paid” (p. 123). Associated with economic drivers are questions of intellectual property rights management; the techniques described by Arms (2000) and Stefik (1999) are essentially focused on the management of intellectual property rights and the integrity of the object (or article). This information was presented above in the discussion on intellectual property.

Related to information security are issues of user confidentiality and privacy, as reported by the American Association for the Advancement of Science (1998). Turner (1998) describes the policies and practices followed at the online *Journal of Electronic Publishing*. This topic—what constitutes appropriate and adequate protection of users of Web-based information—is currently an area of active discussion, as the recent passage of the Electronic Communication Privacy Act, which directly affects publishers of medical information, suggests. Turner’s study did not uncover material unique to users of electronic scientific journals, although it may be

inferred that general practices that emerge in the context of the Web for noncommercial and commercial sites will apply.⁵³

Of greater interest to the research community is how protection of privacy applies to the conduct of research projects as well as to the operation of systems. Yu and Apps (2000, pp. 316–17) describe the approach taken by the SuperJournal project. The project collected sufficient information to link each user with contextual information, but the data were kept strictly confidential through creation of a private file in which the original identifying information was stored. User codes were substituted in subsequent log handling, and access to the identificational file by project staff was limited. Journal names were also masked so that publishers’ business information was protected from their competitors.

To users, information security issues are typically manifested in the form of passwords that control their access to a system. Luther (1997) cautions librarians who acquire e-journal services to identify the provision made for “security (IP addresses or passwords)” to determine compatibility with their libraries’ technical infrastructure. Rusch-Feja and Siebeky (1999a, 1999b) report that comments received during the survey of Max Planck Society researchers reflected users’ desire for unification of access passwords (some systems require multiple passwords if there are several access points, i.e., to more than one system or to more than one journal) and/or better information concerning identification and passwords. One survey respondent noted that electronic journals can only be “reasonable if one can access the journals of different publishers via a central user interface.”

The SuperJournal project (*Summary of SuperJournal Findings: Readers* 1999) required all users to register via a two-step process that required a user name and password that were used once and an e-mail address and personal ID that were entered and then used subsequently. Some SuperJournal users were confused by the two-step process, and follow-up studies showed that, although both users and nonusers appreciated the significance of the password process in authenticating users to the system, they would have preferred that the system recognize them. The most common problem for SuperJournal users was remembering their personal IDs, even though they had selected these themselves; this was also a problem for users of the online

⁵³See, for example, footnote 39 in the Findings section.

Journal of Animal Science (Lewis and Edwards 1998, p. 267).

The views voiced by respondents to the Max Planck Society survey and the SuperJournal project are consistent with more general studies of user behavior with respect to passwords; these document frustration with complex password policies. Adams and Sasse (1999) report findings of a Web-based survey of 139 individuals, approximately half of whom were employees of a technology company and the remainder users in organizations throughout the world. The survey was amplified by follow-up interviews with people in the technology

company and in a construction company. The researchers found that respondents generally ignored or worked around procedures intended to increase security (e.g., violating the sensible stricture not to write down a password) when these requirements seemed to “get in the way.” The respondents also did not appear to understand the reasons for system security requirements—a failing the authors attribute to insufficient communication stemming from an authoritarian approach. System owners tend to guard access to information, including information about security, on the basis of “need to know.” Adams and Sasse argue that more information would in fact induce more responsible user behavior.